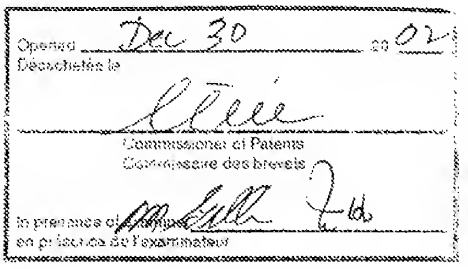
DATE FILED: 05/06/2009
DOCUMENT NO: 48



SUGANO EXHIBIT 1012
FIERS V. SUGANO
INTERFERENCE NO. 105,661

This is EXHIBIT FIERS-31
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

Opposition to European Patent No. 0 041 313 of Biogen, Inc.
Schering Aktiengesellschaft
Appeal File No. T0207/94-334

AFFIDAVIT OF DR. RIK DERYNCK

I, RIK DERYNCK, declare that:

1. I am a citizen of Belgium, and reside at 2 Clark Dr., Apt. 320, San Mateo, CA 94401.

2. I am a Professor in the Department of Growth and Development, the Department of Anatomy, and the Programs of Cell Biology and Developmental Biology, at the University of California, San Francisco, California, U.S.A.

My scientific expertise in the field of Molecular Biology is established by my curriculum vitae, which is attached.

3. In the years 1978-1980, several research groups were engaged in an intense, competitive effort to clone and express the gene for human fibroblast interferon-β (IFN-β). I was a member of one such group, centered at the Laboratory of Molecular Biology, State University of Ghent, in Ghent, Belgium, in the laboratory of Dr. Walter Fiers. This research was sponsored in part by Biogen N.V., now Biogen, Inc. My task was to clone the IFN-β cDNA, and then to insert it in an expression vector to express IFN-β in *E. coli*.

4. After the announcement by Taniguchi in late February 1980 that he had isolated the full size coding sequence of IFN-β, we continued and even increased our efforts to express recombinant IFN-β.

5. The primary goal of this project in our laboratory in expressing the IFN-β cDNA was not to maximize the yield or to achieve commercially or pharmaceutically useful levels of IFN-β in a pure state. These were issues to be later solved by Biogen, or to be the subject of a future project in our laboratory. Rather, our goal was to achieve a finite level of expression of IFN-β, however low, in order to prove the feasibility of the commercial production of the recombinant protein, as well as to be the first group to publish on the recombinant synthesis of IFN-β. We pursued this goal with the methodology available to us at that time, and no new technology was required to achieve the expression of IFN-β, which we accomplished about two months after obtaining possession of the full length cDNA encoding it.

Dr. Rik Derynck

Date 9/26/96

B8/B8CIP

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT INTERFERENCES

WALTER C. FIERS)
)
v.)
)
MICHEL REVEL and) Interference No. 101,096
PIERRE TIOLLAIS)
)
v.)
)
HARUO SUGANO ET AL.)

MOTION OF JUNIOR PARTY WALTER C. FIERS TO AMEND THE ISSUE BY SUBSTITUTION OF PROPOSED COUNT 2 FOR COUNT 1 OR BY ADDITION OF PROPOSED COUNT 2 AND BY ADDITION OF PROPOSED COUNTS 3-9 PURSUANT TO 37 C.F.R. § 1.231(a)(2) AND TO BE ACCORDED THE BENEFIT OF THE FILING DATES OF HIS EARLIER UNITED KINGDOM PATENT APPLICATIONS WITH RESPECT TO THE SUBSTITUTED AND ADDED COUNTS PURSUANT TO 37 C.F.R. § 1.231(a)(4) AND (c)

July 20, 1984

James F. Haley, Jr. (Reg. No. 27,794)
c/o FISH & NEAVE
875 Third Avenue, 29th Floor
New York, New York 10022
Tel.: (212) 715-0600

Attorney for Junior Party
Walter C. Fiers

TABLE OF CONTENTS

TABLE OF AUTHORITIES

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE
BEFORE THE BOARD OF PATENT INTERFERENCES

WALTER C. FIERS	)	
	)	
v.	)	
	)	
MICHEL REVEL and	)	
PIERRE TIOLLAIS	)	Interference No. 101,096
	)	
v.	)	
	)	
HARUO SUGANO ET AL.	)	

MOTION OF JUNIOR PARTY WALTER C. FIERS TO AMEND THE ISSUE BY SUBSTITUTION OF PROPOSED COUNT 2 FOR COUNT 1 OR BY ADDITION OF PROPOSED COUNT 2 AND BY ADDITION OF PROPOSED COUNTS 3-9 PURSUANT TO 37 C.F.R. § 1.231(a)(2) AND TO BE ACCORDED THE BENEFIT OF THE FILING DATES OF HIS EARLIER UNITED KINGDOM PATENT APPLICATIONS WITH RESPECT TO THE SUBSTITUTED AND ADDED COUNTS PURSUANT TO 37 C.F.R. § 1.231(a)(4) AND (c)

Junior Party Walter C. Fiers moves under Rule 231(a)(2) [37 C.F.R. § 1.231(a)(2)] to amend the issue in this interference by substitution of proposed count 2 for Count 1, or by addition of proposed count 2, and by addition of proposed counts 3-9. Fiers also moves to be accorded for the subject matters of proposed counts 2-9 the benefit of the filing dates of his United Kingdom application 8011306, filed April 3, 1980, and his United Kingdom application 8018701, filed June 6, 1980.

Proposed Counts 2-9 are:

Proposed Count 2

A DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon.

Proposed Count 3

A recombinant DNA molecule characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA sequence being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed Count 4

A microbial host transformed with at least one recombinant DNA molecule, said recombinant DNA molecule being characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA sequence being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed Count 5

A human fibroblast β_1 interferon produced by the method of culturing a microbial host transformed with a recombinant DNA molecule characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA sequence being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed Count 6

A method for producing a human fibroblast β_1 interferon comprising the step of culturing a microbial host transformed with at least one recombinant DNA molecule, said recombinant DNA molecule being characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed Count 7

A method for selecting a DNA sequence which codes for a human fibroblast β_1 interferon from a group of DNA sequences comprising the step of determining which of said DNA sequences hybridizes to a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon.

Proposed Count 8

A composition for treating human viruses, treating human cancers or tumors, or useful in immunomodulation, comprising a pharmaceutically effective amount of at least one human fibroblast β_1 interferon produced by the method of culturing a microbial host transformed with a recombinant DNA molecule characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA sequence being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed Count 9

A method for treating human viruses, treating human cancers or tumors, or useful in immunomodulation, comprising the step of administering to humans in a pharmaceutically acceptable manner a pharmaceutically effective amount of a composition comprising a human fibroblast β_1 interferon produced by the method of culturing a microbial host transformed with a recombinant DNA molecule characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, said DNA sequence being operatively linked to an expression control sequence in the recombinant DNA molecule.

Proposed count 2 should be substituted for Count 1, or in the alternative added to this interference, because proposed count 2 is preferable to Count 1. It omits from Count 1 the term "polypeptide" which is unnecessary and may unreasonably complicate the proofs with respect to Count 1. Proposed count 2 also more correctly defines the DNA sequences of this interference in terms of a family of human fibroblast β_1 interferons.

Proposed count 2 is patentable to junior party Fiers and appears in his application.* It is "patentable" to Revel and Tiollais and to Sugano et al. and "supported" in their applications for the same alleged reasons that Count 1 is patentable to and supported by them.**

---

* Fiers has concurrently herewith amended his United States patent application 250,609 to add, as claim 32, a claim that corresponds in all respects to proposed count 2. Fiers has appended hereto as Exhibit A a copy of that Amendment.

** Fiers has moved concurrently herewith to dissolve this interference on the grounds that Count 1 is not patentable to Revel and Tiollais and not supported in their United States applications. See Motion Of Junior Party Walter C. Fiers To Dissolve This Interference Pursuant To 37 C.F.R. § 1.231(a)(1) As To Revel And Tiollais And To Deny To Revel And Tiollais The Benefit Of The Filing Date Of Their United States Patent Application 208,925 ("Fiers Motion To Dissolve-Revel And Tiollais).

Fiers has also moved to dissolve this interference because Count 1 is not patentable to Sugano et al. See Motion Of Junior Party Walter C. Fiers To Dissolve This Interference Pursuant To 37 C.F.R. § 1.231(a)(1) As To Sugano et al ("Fiers Motion To Dissolve-Sugano et al.").

Proposed counts 3-9 should be added to this interference because they are necessary to resolve all of the potential priority issues between the applications in interference. They involve different potential proofs. And they are patentably distinct from each other and the subject matter of Count 1 and proposed count 2. Each of proposed counts 3-9 is patentable to junior party Fiers and appears in his application.* The other parties to this interference may also believe, and the Patent Examiner may find, that one or more of proposed counts 3-9 are patentable to them (which they are not), or supported by their applications (which they are not).

By this motion, junior party Fiers provides to Sugano et al. and to Revel and Tiollais the opportunity to attempt to demonstrate that one or more of proposed counts 3-9 are patentable to them and that their applications support one or more of those counts (which Fiers believes they cannot do). This motion also puts Sugano et al. and Revel and Tiollais on notice that junior party Fiers believes that the subject matters of proposed counts 3-9 are part of the invention that he has described and claimed in his application and that he intends to pursue the grant to him of United States patents describing and claiming the subject matters of proposed counts 3-9.

---

* Fiers has concurrently herewith amended his United States patent application 250,609 to add, as claims 33-39, claims that correspond in all respects to proposed counts 3-9. Fiers has appended hereto as Exhibit A a copy of that Amendment.

I. PROPOSED COUNT 2 SHOULD BE SUBSTITUTED FOR COUNT 1 OR IN THE ALTERNATIVE ADDED TO THIS INTERFERENCE

A. The Substitution Or Addition Of Proposed Count 2 To This Interference Is Necessary And Preferable

Count 1 did not originally appear in any of the United States patent applications now in interference. The Examiner suggested Count 1 for purposes of interference in Revel and Tiollais application 425,934 and in Sugano et al. application 201,359 on June 28, 1983. The Examiner subsequently suggested Count 1 for purposes of interference in Fiers application 250,609 on September 26, 1983.

Because Count 1 did not originally appear in any of the applications in interference, it contains a term -- "polypeptide" -- which is unnecessary and may unreasonably complicate the parties' proofs with respect to Count 1. Count 1 is also ambiguous with respect to whether or not it includes, as it should, DNA sequences that code for the family of human fibroblast β_1 interferons. Proposed count 2 avoids each of these potential difficulties. Accordingly, it should be substituted for Count 1, or in the alternative added to this interference.

Proposed count 2 omits the unnecessary term "polypeptide" from Count 1. Count 1 defines a DNA sequence that codes for a product of specified activity -- a human fibroblast β_1 interferon. For purposes of this interference no further characterization of that interferon activity is required. The omission of the term "polypeptide" from Count 1 also avoids potential objections to the parties' disclosures and proofs that the human fibroblast β_1 interferon coded for by the claimed DNA sequence is a polypeptide. For example, the term "polypeptide

could arguably require an actual determination of the complete amino acid sequence of a human fibroblast β_1 interferon and, perhaps, a demonstration that the bonds between the various amino acids are peptide bonds. It could also arguably require a demonstration that the product is sensitive to proteases. Because the term "polypeptide" is unnecessary, a requirement for these proofs is also not appropriate.

Proposed count 2 also avoids the ambiguity of Count 1 by reciting that the DNA sequence codes for "a" human fibroblast β_1 interferon. Accordingly, proposed count 2 acknowledges that the recited human fibroblast β_1 interferon is actually a family of products.* It also avoids difficult to resolve disputes of whether one or more amino acid additions or deletions from native human fibroblast β_1 interferon in an interferon coded for by the recited DNA sequence take that product out of the scope of Count 1. Accordingly, proposed count 2's recitation of a genus of human fibroblast β_1 interferons is also necessary and preferable for contesting the priority issues in interference.

B. Proposed Count 2 Is Patentable To Junior Party Fiers And Appears In His Application

Proposed count 2 is patentable to junior party Fiers and appears in Fiers United States application 250,609.**

---

* Fiers has demonstrated that several different products have the activity of a human fibroblast β_1 interferon. For example, Fiers produced "mature" fibroblast β_1 interferon as the active product of G-pPLa-HFIF-67-12ΔM1 and G-pHLa-HFIF-67-12Δ19 BX-2 [page 88, lines 19-22]. Fiers also produced possible fusion products that had IFN-β_1 activity from G-pPLa-HFIF-67-12Δ19 and pPLc-HFIF-67-8 [page 88, lines 22-31].

** Fiers has concurrently herewith amended his United States patent application 250,609 to add, as claim 32, a claim that corresponds in all respects to proposed count 2. Fiers has appended hereto as Exhibit A a copy of that Amendment.

Proposed count 2 is specifically supported by Fiers application 250,609 for the same reasons as Count 1 is supported by that application. For example, Fiers described a DNA sequence for the "coding strand of IFN-β DNA" [page 48, line 36-page 49, line 2; Figure 4] and disclosed that the DNA sequence was capable of directing the production of "polypeptides displaying an immunological or biological activity of HuIFN-β" [page 12, lines 22-23; page 77, line 31-page 88, line 35]. Fiers also identified and deposited in recognized culture collections before the filing date of that application specific DNA sequences that coded for and produced a human fibroblast β_1 interferon. For example, Fiers identified and deposited E.coli HB101(G-pBR322(Pst)/HFIF3, 6 and 7) that contained DNA sequences from which a DNA sequence that codes for a human fibroblast β_1 interferon is produced [page 94, lines 13-22]. Fiers also described using that DNA sequence and its derivatives to produce various human fibroblast β_1 interferons [e.g., page 77, line 31-page 88, line 35] and Fiers deposited in recognized culture collections specific DNA sequences that produced those β_1 interferons [page 94, line 23-page 95, line 2].*

Proposed count 2 is patentable to Fiers for the same reasons that Count 1 is patentable to him. There is no document of record in Fiers United States application 250,609, or in any of the other applications in interference, having an effective date before April 3, 1980 or June 6, 1980 (the filing dates of the Fiers United Kingdom patent applications 8011306 and 8018701 to which he is entitled for the subject

---

* For example, Fiers produced "mature" fibroblast β_1 interferon as the active product of G-pPLa-HFIF-67-12ΔM1 (ATCC 31824) and G-pPLa-HFIF-67-12Δ19 BX-2 [page 88, lines 19-22]. Fiers also produced possible fusion products that had fibroblast β_1 interferon activity from G-pPLa-HFIF-67-12Δ19 (DSM 1853) and pPLc-HFIF-67-8 (DSM 1854) [page 88, lines 22-31].

matter of Count 1) that describes or suggests that subject matter. See Motion Of Junior Party Walter C. Fiers To Be Accorded The Benefit Of The Filing Dates Of His Earlier United Kingdom Patent Applications Pursuant To 37 C.F.R. §§ 1.224 and 1.231(a)(4) And To Shift The Burden Of Proof Under 37 C.F.R. § 1.257(a) ("Fiers Motion-Count 1"). The deletion of the term "polypeptide" in proposed count 2 and its avoidance of the former "ambiguity" of Count 1 does not affect the novelty and unobviousness of the subject matter of that count to Fiers.

C. Proposed Count 2 And The Sugano et al. And Revel And Tiollais United States Applications

Proposed count 2 is "supported" in the Sugano et al. United States application for the same alleged reasons as Count 1. Sugano et al. United States application 201,359 refers to "a DNA which codes for a human fibroblast interferon polypeptide" and recites in Table 5 a base sequence for that DNA [page 8, lines 12-13]. It also recites that a DNA sequence (T_pIF-319-13) was deposited on September 16, 1980 in a culture collection. However, proposed count 2 is not patentable to Sugano et al. for the same reasons that Count 1 is not patentable to them.* Should Sugano et al. be successful in demonstrating the patentability to themselves of Count 1 (which they should not be able to do), proposed count 2 should be substituted for Count 1, or in the alternative added to this interference.

Proposed count 2 is not "supported" in the Revel and Tiollais United States applications and not "patentable to" them for the same reasons that Count 1 is not supported

---

* Fiers has moved concurrently herewith to dissolve this interference because Count 1 is not patentable to Sugano et al. See Fiers Motion To Dissolve-Sugano et al.

by or patentable to them.* However, should Revel and Tiollais be successful in demonstrating that their United States applications support Count 1 and that Count 1 is patentable to them (which they should not be able to do), proposed count 2 should be substituted for Count 1, or in the alternative added to this interference.

II. PROPOSED COUNTS 3-9 SHOULD BE ADDED TO THIS INTERFERENCE

A. The Addition Of Proposed Counts 3-9 To This Interference Is Necessary And Preferable

The addition of proposed counts 3-9 to this interference is necessary and preferable in order to resolve all of the potential priority issues between the applications in interference in a single interference proceeding.

Each of the United States applications in interference originally recited various groups of claims. During their initial prosecutions, the respective Patent Examiners made multiple restriction requirements in each application among these groups of claims [Sugano et al. application 201,359, Paper No. 3, November 3, 1981; Revel and Tiollais application 208,925, Paper No. 3, November 17, 1981; Fiers application 250,609, Examiner's Action, April 15, 1982]. For example, the Examiners divided the applications among claims drawn to DNA and claims drawn to plasmids and hosts [Sugano et al., Paper No. 3, p. 2]; among claims drawn to DNA, claims drawn to methods of detecting DNA, claims drawn to recombinant DNA molecules, and claims drawn to interferons [Revel and

---

* Fiers has moved concurrently herewith to dissolve this interference with respect to Revel and Tiollais on both of those grounds. See Fiers Motion To Dissolve-Revel and Tiollais.

Tiollais, Paper No. 3, p. 2]; and among claims drawn to DNA, claims drawn to interferons, claims drawn to methods of producing interferons, claims drawn to methods of detecting DNA and claims drawn to methods of treating viruses and cancers [Fiers, Examiner's Action, p. 2]. In support of these restriction requirements, the Examiners contended that the various groups of claims were "distinct" and could support "separate patents".

In response to these restriction requirements, each of the applicants elected claims for initial prosecution and withdrew the non-elected claims from prosecution. Accordingly, each of the applications has claims that are not now in interference and which the Examiners have held to be "distinct" and able to support "separate patents". It is these groups of claims that Fiers has moved to add to this interference.*

Proposed counts 3-9 are patentably distinct from each other and from Count 1 and proposed count 2 [See United States Patent and Trademark Office, Manual of Patenting Examining Procedure (5th ed. August 1983) § 1105.03]. They also represent subject matter on which different proofs of "invention" and "priority" are likely to be offered by each of the parties. For example, the recombinant DNA molecule of proposed count 3 will require different proofs from the DNA of Count 1 because the recombinant DNA molecule must be capable of expressing a human fibroblast β_1 interferon. And the DNA selection method of proposed count 7 will require different proofs from the interferon of proposed count 5.

---

* For example, the human fibroblast β_1 interferon of proposed count 5 was in Fiers Group II, the method of proposed count 6 in Fiers Group III, the method of proposed count 7 in Fiers Group IV and the method of proposed count 9 in Fiers Group V.

Plainly, a single interference should resolve all of the potential priority issues between these applications. Judicial economy demands nothing less. So does the possibility of interference estoppel that may affect junior or losing parties to an interference [e.g., Stoudt v. Guggenheim, 651 F.2d 760, 210 USPQ 359 (CCPA 1981); In re Hoover Co., 134 F.2d 624, 57 USPQ 111 (CCPA 1943); In re Allsop, 26 F.2d 559 (D.C.Cir. 1928); Blackford v. Wilder, 28 App. D.C. 535, 1907 C.D. 491, 127 O.G. 1255 (1907); Ex parte Miller, 124 USPQ 419 (PO Bd App 1959); Ex parte Voris, 92 USPQ 47 (PO Bd App 1951); 37 C.F.R. § 1.257 (1983)]. Therefore, proposed counts 3-9 should be added here. However, they should not be added to this interference unless, and until, the Patent Examiner finds that the proposed counts are patentable to either Sugano et al. or Revel and Tiollais (which they are not) and are supported by their applications (which they are not) [infra, pp. 13-15].

B. Proposed Counts 3-9 Are Patentable To Junior Party Fiers And Appear In His Application

Proposed counts 3-9 are patentable to junior party Fiers and appear in his United States patent application 250,609.*

Proposed counts 3-9 are specifically supported by Fiers application 250,609. For example, Fiers described a DNA sequence for the "coding strand of IFN-β DNA" [page 48, line 36-page 49, line 2; Figure 4] and disclosed that the DNA sequence was capable when operatively linked to an expression control sequence in a recombinant DNA molecule of directing

---

* Fiers has concurrently herewith amended his United States patent application 250,609 to add, as claims 33-39, claims that correspond in all respects to proposed counts 3-9. Fiers has appended hereto as Exhibit A a copy of that Amendment.

the production of "polypeptides displaying an immunological or biological activity of HuIFN-β" in microbial hosts transformed with that recombinant DNA molecule [page 12, lines 22-23; page 77, line 31-page 88, line 35]. Fiers also identified and deposited in recognized culture collections recombinant DNA molecules characterized by DNA sequences that code for a human fibroblast β_1 interferon and are operatively linked to expression control sequences (proposed count 3), and microbial hosts transformed with those recombinant DNA molecules (proposed count 4) that produced a human fibroblast β_1 interferon after fermentation (proposed counts 5 and 6) [page 94, line 22-page 95, line 2].

Fiers also disclosed using his DNA sequences that coded for a human fibroblast β_1 interferon to select by hybridization other DNA sequences that coded for a human fibroblast β_1 interferon (proposed count 7) [e.g., page 44, lines 26-page 47, line 15; page 92, line 31-page 94, line 12]. And Fiers disclosed using his human fibroblast β_1 interferons in compositions (proposed count 8) and methods (proposed count 9) for treating human viruses, treating human cancers or tumors, or in immunomodulation [e.g., page 4, line 11-page 7, line 15; page 12, lines 29-35].

Finally, Fiers originally claimed the subject matter of proposed counts 3-9 in his United States application. Compare proposed count 3 with Fiers claims 7-10; proposed count 4 with Fiers claim 11 (as it depends from claims 7-10) and claims 12-15; proposed count 5 with Fiers claims 16 and claims 18-19 (as they depend from claim 16); proposed count 6 with Fiers claim 25; proposed claim 7 with Fiers claims 26-27; proposed count 8 with Fiers claim 28 and; proposed count 9 with Fiers claim 29.

As we have demonstrated, proposed counts 3-9 are supported in the Fiers United States application. They are also patentable to him. None of the documents of record in Fiers United States application 250,609, or in any of the other United States applications in interference, described or suggested the production of a human fibroblast β_1 interferon in a microbial host before the dates to which Fiers is entitled for proposed counts 3-9 [infra, pp. 15-18]. None suggested or described any recombinant DNA molecule characterized by a DNA sequence which consists essentially of a DNA which codes for a human fibroblast β_1 interferon, which DNA is operatively linked to an expression control sequence in the recombinant DNA molecule (proposed count 3) or any microbial host transformed with that recombinant DNA molecule (proposed count 4). None suggested or described using those microbial hosts to produce a human fibroblast β_1 interferon by fermentation (proposed count 6), the interferon produced by that method (proposed count 5), or methods or compositions employing that interferon against viruses, cancer and tumors, or in immunomodulation (proposed counts 8-9). Furthermore, none suggested or described using that IFN-β_1 DNA sequence to select other DNA sequences that also code for a human fibroblast β_1 interferon by hybridization (proposed count 7).

C. Proposed Counts 3-9 And The Sugano et al. And Revel And Tiollais United States Applications

Neither Sugano et al. United States patent application 201,359, nor Revel and Tiollais United States patent application 208,925, supports the subject matter of proposed counts 3-9.

Neither Sugano et al. nor Revel and Tiollais described any recombinant DNA molecule characterized by a DNA sequence

consisting essentially of a DNA coding for a human fibroblast β_1 interferon which is operatively linked to an expression control sequence in the recombinant DNA molecule (proposed count 3) or any microbial host transformed with that recombinant DNA molecule (proposed count 4). Neither described using a microbial host to produce a human fibroblast β_1 interferon by fermentation (proposed count 6), any interferon produced by that method (proposed count 5), or any method or composition for employing that interferon against viruses, cancers or tumors, or in immunodulation (proposed counts 8-9). And neither described using that DNA sequence to select other DNA sequences that code for a human fibroblast β_1 interferon (proposed count 7). Furthermore, neither Sugano et al. nor Revel and Tiollais deposited in a culture collection any recombinant DNA molecule or microbial host that produced a human fibroblast β_1 interferon. Accordingly, neither is entitled to make claims corresponding to proposed counts 3-9.

Sugano et al. implicitly concedes that they have no right to make claims corresponding to proposed counts 3-9 by not claiming that subject matter in their United States application 201,359. Revel and Tiollais United States application 208,925 does "claim" a process for "engineering a bacterial strain to produce interferon polypeptide" using a "suitable vector-carrier" (Revel and Tiollais claims 9-11) and the "interferon" produced by that process (Revel and Tiollais claims 24, 43-44). However, neither of those classes of claims is supported in any way in Revel and Tiollais' United States applications. See Fiers Motion To Dissolve-Revel And Tiollais. Accordingly, neither Sugano et al. nor Revel and Tiollais are entitled to contest the priority of the subject matter of proposed counts 3-9.

Although junior party Fiers does not believe that Sugano et al. or Revel and Tiollais can make any claim corresponding to proposed counts 3-9, Fiers has by this motion given to each of them the opportunity to demonstrate their "support" for the subject matters of those proposed counts and the "patentability" of those proposed counts to themselves. Fiers has also put each of them on notice that he intends to pursue the grant of United States patents describing and claiming the subject matters of proposed counts 3-9.

III. FIERS SHOULD BE ACCORDED THE BENEFIT OF THE FILING DATES OF HIS EARLIER UNITED KINGDOM PATENT APPLICATIONS FOR PROPOSED COUNTS 2-9 AND SHOULD BE SENIOR PARTY WITH RESPECT TO THOSE PROPOSED COUNTS

Junior party Fiers moves under Rule 224 and 231(a)(4) [37 C.F.R. §§ 1.224 and 1.231(a)(4)] to be accorded the benefit of the filing dates of his earlier United Kingdom patent applications for the subject matter of proposed counts 2-9 and to be accorded senior party status with respect to those proposed counts under Rule 257(a) [37 C.F.R. § 1.257(a)].

A. Fiers Is Entitled To The Benefit Of The April 3, 1980 Filing Date Of His United Kingdom Patent Application 8011306 For The Subject Matters Of Proposed Counts 2 and 7

Fiers United Kingdom patent application 8011306 describes and enables the subject matter of proposed count 2 for the same reasons that it describes and enables the subject matter of Count 1.* See Fiers Motion-Count 1, pp. 2-6. The differences between Count 1 and proposed count 2 do not change in any way Fiers' specific support for that subject matter.

---

* A certified copy of that application was of record in Fiers United States application 250,609 at the time this interference was declared. Accordingly, it is not submitted here. See 37 C.F.R. § 1.224.

Fiers United Kingdom patent application 8011306 also supports proposed count 7. It described and claimed a method for using a DNA sequence that consisted essentially of DNA coding for a human fibroblast β_1 interferon to select other DNA sequences coding for a human fibroblast β_1 interferon by hybridization (proposed count 7). See, e.g., page 40, line 1-page 42, line 14; claims 12-14.

Fiers previously claimed priority from that United Kingdom application in his United States application 250,609 and made a certified copy of it of record there. Nothing else needs to be done for Fiers to be entitled to the filing date of that application for the subject matters of proposed counts 2 and 7.

B. Fiers Is Entitled To The Benefit Of The June 6, 1980 Filing Date Of His United Kingdom Patent Application 8018701 For The Subject Matters Of Proposed Counts 2 and 7

Fiers United Kingdom patent application 8018701 describes and enables the subject matter of proposed count 2 for the same reasons that it describes and enables the subject matter of Count 1.* See Fiers Motion-Count 1, pp. 6-8.

Fiers United Kingdom patent application 8018701 also described and claimed a method for using a DNA sequence that consisted essentially of DNA coding for a human fibroblast β_1 interferon to select other DNA sequences coding for a human fibroblast β_1 interferon by hybridization (proposed count 7). See, e.g., page 42, line 1-page 44, line 18; claims 12-14.

---

* A certified copy of that application was of record in Fiers United States application 250,609 at the time this interference was declared. Accordingly, it is not submitted here. See 37 C.F.R. § 1.224.

Fiers previously claimed priority from that United Kingdom application in his United States application 250,609 and made a certified copy of it of record there. Nothing else needs to be done for Fiers to be entitled to the filing date of that application for the subject matters of proposed counts 2 and 7.

C. Fiers Is Entitled To The Benefit Of The June 6, 1980 Filing Date Of His United Kingdom Patent Application 8018701 For The Subject Matters Of Proposed Counts 3-6 And 8-9

Fiers United Kingdom patent application 8018701 describes and enables the subject matters of proposed counts 3-6 and 8-9. It also sets forth the best mode contemplated by Fiers for practicing those subject matters. Accordingly, Fiers should be accorded for the subject matters of proposed counts 3-6 and 8-9 the benefit of his June 6, 1980 filing date of United Kingdom patent application 8018701.*

United Kingdom patent application 8018701, filed June 6, 1980, described recombinant DNA molecules characterized by a DNA sequence which consisted essentially of a DNA that codes for a human fibroblast β_1 interferon and which was operatively linked to an expression control sequence in those recombinant DNA molecules (proposed count 3) [page 54, line 1-page 60, line 11]. It described microbial hosts transformed with those recombinant DNA molecules (proposed count 4) that produced a human fibroblast β_1 interferon after fermentation (proposed counts 5 and 6) [page 61, line 1-page 76, line 26]. The application also described using the produced human fibroblast β_1

---

* Fiers has not attached to this motion a copy of his United Kingdom application 8018701. A certified copy of the application was of record in Fiers United States application 250,609 at the time this interference was declared. See 37 C.F.R. § 1.224.

interferon in compositions (proposed count 8) and methods (proposed count 9) for treating human viruses and cancers or tumors [page 5, line 3-page 7, line 25; page 11, lines 30-35].

Fiers United Kingdom application 8018701 also recited the deposit in a recognized culture collection of recombinant DNA molecules characterized by a DNA sequence that consisted essentially of a DNA that coded for a human fibroblast β_1 interferon which was operatively linked to an expression control sequence in that recombinant DNA molecule, and microbial hosts transformed with those recombinant DNA molecules that produced a human fibroblast β_1 interferon upon fermentation [page 81, lines 10-19].

Finally, Fiers United Kingdom application 8018701 originally claimed the subject matter of proposed counts 3-6 and 8-9. Compare, e.g., proposed count 3 with Fiers claims 22, 25 and 26; proposed count 4 with Fiers claims 27-28 (as they depend from claims 22, 25 and 26) and claim 29; proposed count 5 with Fiers claim 32 (as it depends from claims 22, 25 and 26); proposed count 6 with Fiers claims 39 and 40; proposed count 8 with Fiers claims 35 and 37 (as they depend from claims 22, 25 and 26) and claim 36 (as it depends from claim 32); and proposed count 9 with Fiers claims 41-43 (as they depend from claims 22, 25 and 26) and claims 42 and 44 (as they depend from claim 32).

For all of those reasons, Fiers is entitled to the benefit of the June 6, 1980 filing date of his United Kingdom patent application 8018701 for the subject matters of proposed counts 3-6 and 8-9.

IV. CONCLUSION

Proposed count 2 should be substituted for Count 1, or in the alternative added to this interference. It is pre-

ferable to Count 1 and necessary to avoid the ambiguity of Count 1 and the possible objections to the parties' applications and proofs with respect to Count 1. If proposed count 2 is added to this interference, or substituted for Count 1, Fiers is entitled to the April 3, 1980 filing date of his United Kingdom patent application 8011306 and to the June 6, 1980 filing date of his United Kingdom patent application 8018701 for the subject matter of that proposed count. Accordingly, Fiers is senior party with respect to proposed count 2.

Proposed counts 3-9 should be added to this interference if, and only if, the Examiner finds that Sugano et al. or Revel and Tiollais support them (which they do not) and that the proposed counts are patentable to them (which they are not).

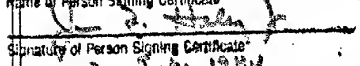
If any of the proposed counts 3-9 is added to this interference, Fiers is entitled to the filing dates of his earlier United Kingdom patent applications for their subject matters. He is entitled to the April 3, 1980 filing date of his United Kingdom patent application 8011306 and to the June 6, 1980 filing date of his United Kingdom patent application 8018701 for the subject matter of proposed count 7. He is entitled to the June 6, 1980 filing date of his United Kingdom patent application 8018701 for the subject matters of proposed counts 3-6 and 8-9. Accordingly, Fiers is senior party with respect to each of proposed counts 3-9.

Respectfully submitted,

James F. Haley, Jr.
Attorney of Record
Registration No. 27,794
c/o FISH & NEAVE
875 Third Avenue
29th Floor
New York, New York 10022
Tel.: (212) 715-0600

I hereby certify that this correspondence is being deposited with the United States Postal Service as first class mail in an envelope addressed to: Commissioner of Patents and Trademarks, Washington, D.C. 20231, on 2nd July 1984

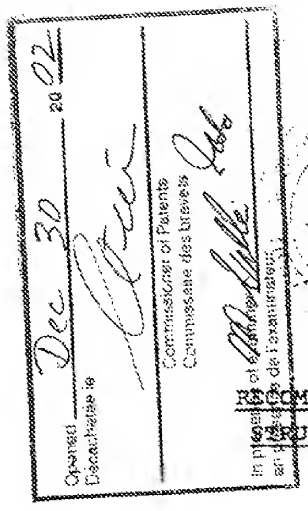
James F. Haley Jr.
Name of Person Signing Certificate



Signature of Person Signing Certificate

2nd July 1984
Date of Signature

FIF
88
DRAFT



RECOMBINANT DNA MOLECULES AND THEIR USE IN PRODUCING STRUCTURAL GENES FOR HUMAN FIBROBLAST INTERFERON

TECHNICAL FIELD OF INVENTION

This invention relates to recombinant DNA molecules and their use in producing structural genes for human fibroblast interferon. The recombinant DNA molecules disclosed herein are characterized by DNA sequences that code for polypeptides whose amino acid sequence and composition are substantially consistent with human fibroblast interferon.

BACKGROUND ART

Two classes of interferons ("IF") are known to exist. Interferons of Class I are small, acid stable (glyco)-proteins that render cells resistant to viral infection (A. Isaacs and J. Lindenmann, "Virus Interference I. The Interferon", Proc. Royal Soc. Ser. B., 147, pp. 258-67 (1957) and W.E. Stewart, II, The Interferon System, Springer-Verlag (1979) (hereinafter "The Interferon System")). Class II IFs are acid labile. At present, they are poorly characterized. Although to some extent cell specific (The Interferon System, pp. 135-45), IFs are not virus specific. Instead, IFs protect cells against a wide spectrum of viruses.

Two antigenically distinct species of Class I human interferon ("HIF") are known to exhibit IF activity. One IF species, fibroblast interferon ("F IF"), is produced upon appropriate induction in diploid fibroblast cells. Another IF species, leukocyte interferon ("Le IF") is produced together with minor amounts of F IF upon appropriate induction in human leukocyte and lympho-

blastoid cells. Both are heterogeneous in regard to size, presumably because of the carbohydrate moiety. F IF has been extensively purified and characterized (E. Knight, Jr., "Interferon: Purification and Initial Characterization From Human Diploid Cells", Proc. Natl. Acad. Sci. USA, 73, pp. 520-23 (1976)). It is a glyco-protein of about 20,000 molecular weight (M. Wiranowska-Stewart et al., "Contributions Of Carbohydrate Moieties To The Physical And Biological Properties Of Human Leukocyte, Lymphoblastoid And Fibroblast Interferons", Abst. Ann. Meeting Amer. Soc. Microbiol., p. 246 (1978)).

Its amino acid composition has been determined (E. Knight, Jr. et al., "Human Fibroblast Interferon: Amino Acid Analysis and Amino-Terminal Amino Acid Sequence", Science, 207, pp. 525-26 (1980)). Elucidation of its amino acid sequence is in progress. To date, the amino acid sequence of the NH_2 terminus of the mature protein has been reported for the first 13 amino acid residues: Met-Ser-Tyr-Asn-Leu-Leu-Gly-Phe-Leu-Gln-Arg-Ser-Ser... (E. Knight, Jr. et al., supra). Two distinct genes, one located on chromosome 2, the other on chromosome 5, have been reported to code for F IF (D.L. Slate and F.H. Ruddle, "Fibroblast Interferon In Man Is Coded By Two Loci On Separate Chromosomes", Cell, 16, pp. 171-80 (1979)).

Le IF has likewise been purified and characterized. Two components have been described, one of 21000 to 22000 and the other of 15000 to 18000 molecular weight (K.C. Zoon, et al., "Purification And Partial Characterization Of Human Lymphoblastoid Interferon", Proc. Natl. Acad. Sci. USA, 76, pp. 5601-605 (1979)). A portion of the amino acid sequence of Le IF has also been determined, i.e., 20 amino acids from the amino terminus of the mature protein (K.C. Zoon et al., "Amino-Terminal Sequence Of The Major Component Of Human Lymphoblastoid Interferon", Science, 207, pp. 527-28 (1980)). A comparison of the initial amino acid sequence of F IF and Le IF reveals no detectable homology within the first 13 amino acids. The total amino acid compositions of the two species are also distinct. In addition, degradation of the sugar residues of the two species by periodate indicates that the carbohydrate structure of the two IFs is different (M. Wiranowska-Stewart et al., supra).

The two species of HIF have a number of different properties. For example, anti-human Le IF antibodies are less efficient

against F IF and anti-sera to human F IF have no activity against human Le IF (The Interferon System, p. 151) and Le IF displays a high degree of activity in cell cultures of bovine, feline or porcine origin whereas F IF is hardly active in those cells but has been reported to be active in rat cells (P. Duc-Goiran et al., "Studies On Virus-Induced Interferons Produced By the Human Amniotic Membrane And White Blood Cells", Arch. Ges. Virus Forsch., 34, pp. 232-43 (1971)). In addition, the two IFs result from different mRNA species (and therefore from presumable different structural genes) that code for polypeptides of different primary sequence (R.L. Cavalieri et al., "Synthesis of Human Interferon by Xenopus laevis Oocytes: Two Structural Genes for Interferon in Human Cells", Proc. Natl. Acad. Sci. USA, 74, pp. 3287-91 (1977)).

Although both Le and F IFs occur in a glycosylated form, removal of the carbohydrate moiety (P.J. Bridgen et al., "Human Lymphoblastoid Interferon", J. Biol. Chem., 252, pp. 6585-87 (1977)) or synthesis of IF in the presence of inhibitors which preclude glycosylation (W.E. Stewart, II et al., "Effect of Glycosylation Inhibitors On The Production And Properties Of Human Leukocyte Interferon", Virology, 97, pp. 473-76 (1979); J. Fujisawa et al., "Nonglycosylated Mouse L Cell Interferon Produced By The Action Of Tunicamycin", J. Biol. Chem., 253, pp. 8677-79 (1978); E.A. Havell et al., "Altered Molecular Species Of Human Interferon Produced In The Presence Of Inhibitors Of Glycosylation", J. Biol. Chem., 252, pp. 4425-27 (1977); The Interferon System, p. 181) yields a smaller form of IF which still retains most or all of its IF activity.

Both F IF and Le IF may, like many human proteins, be polymorphic. Therefore, cells of particular individuals may produce IF species within each of the more general F IF and Le IF classes which are physiologically similar but structurally slightly different than the class of which it is a part. Therefore, while the protein structure of the F IF or Le IF may be generally well-defined, particular individuals may produce IFs that are slight variations thereof.

IF is usually not detectable in normal or healthy cells (The Interferon System, pp. 55-57). Instead, the protein is produced as a result of the cell's exposure to an IF inducer. IF inducers are usually viruses but may also be non-viral in character,

such as natural or synthetic double-stranded RNA, intracellular microbes, microbial products and various chemical agents. Numerous attempts have been made to take advantage of these non-viral inducers to render human cells resistant to viral infection (S. Baron and F. Dianzani (eds.), Texas Reports On Biology And Medicine, 35 ("Texas Reports"), pp. 528-40 (1977)). These attempts have not been very successful. Instead, use of exogenous IF itself is now preferred.

As an antiviral agent, HIF has been used to treat the following: respiratory infections (Texas Reports, pp. 486-96); herpes simplex keratitis (Texas Reports, pp. 497-500; R. Sundmacher, "Exogenous Interferon in Eye Diseases", International Virology IV, The Hague, Abstract nr. W2/11, p. 99 (1978)); acute hemorrhagic conjunctivitis (Texas Reports, pp. 501-10); adenovirus keratoconjunctivitis (A. Romano et al., ISM Memo I-A8131 (October, 1979)); varicella zoster (Texas Reports, pp. 511-15); cytomegalovirus infection (Texas Reports, pp. 523-27); and hepatitis B (Texas Reports, pp. 516-22). See also The Interferon System, pp. 307-19. In these treatments F IF and Le IF may display different dose/response curves. However, large-scale use of IF as an antiviral agent requires larger amounts of HIF than heretofore have been available.

IF has other effects in addition to its antiviral action. For example, it antagonizes the effect of colony stimulating factor, inhibits the growth of hemopoietic colony-forming cells and interfers with the normal differentiation of granulocyte and macrophage precursors (Texas Reports, pp. 343-49). It also inhibits erythroid differentiation in DMSO-treated Friend leukemia cells (Texas Reports, pp. 420-28). Some cell lines may be considerably more sensitive to F IF than to Le IF in these regards (S. Einhorn and H. Strander, "Is Interferon Tissue-Specific? - Effect Of Human Leukocyte And Fibroblast Interferons On The Growth Of Lymphoblastoid And Osteosarcoma Cell Lines", J. Gen. Virol., 35, pp. 573-77 (1977); T. Kuwata et al., "Comparison Of The Suppression Of Cell And Virus Growth In Transformed Human Cells By Leukocyte And Fibroblast Interferon", J. Gen. Virol., 43, pp. 435-39 (1979)).

IF may also play a role in regulation of the immune response. For example, depending upon the dose and time of application in relation to antigen, IF can be both immunopotentiating and immunosuppressive in vivo and in vitro (Texas Reports, pp. 357-69). In

In addition, specifically sensitized lymphocytes have been observed to produce IF after contact with antigen. Such antigen-induced IF could therefore be a regulator of the immune response, affecting both circulating antigen levels and expression of cellular immunity (Texas Reports, pp. 370-74). IF is also known to enhance the activity of killer lymphocytes and antibody-dependent cell-mediated cytotoxicity (R.R. Herberman et al., "Augmentation By Interferon Of Human Natural And Antibody-Dependent Cell-Mediated Cytotoxicity", Nature, 277, pp. 221-23 (1979); P. Beverley and D. Knight, "Killing Comes Naturally", Nature, 278, pp. 119-20 (1979); Texas Reports, pp. 375-80; M. Lucero et al., "Induction And Kinetics Of Natural Killer Cells in Humans Following Interferon Therapy", Nature, 282, pp. 417-19 (1979); S. Einhorn et al.,

Acta Med. Scand., 20, pp. 477-83 (1978)). Both may be directly or indirectly involved in the immunological attack on tumor cells.

Therefore, in addition to its use as a human antiviral agent, HIF has potential application in antitumor and anticancer therapy (The Interferon System, pp. 319-21 and 394-99). It is now known that IFs affect the growth of many classes of tumors in many animals (The Interferon System, pp. 292-304). They, like other anti-tumor agents, seem most effective when directed against small tumors. The antitumor effects of animal IF are dependent on dosage and time but have been demonstrated at concentrations below toxic levels. Accordingly, numerous investigations and clinical trials have been and continue to be conducted into the antitumor and anticancer properties of HIFs. These include treatment of several malignant diseases such as osteosarcoma, acute myeloid leukemia, multiple myeloma and Hodgkin's disease (Texas Reports, pp. 429-35). In addition, F IF has recently been shown to cause local tumor regression when injected into subcutaneous tumoral nodules in melanoma and breast carcinoma-affected patients (T. Nemoto et al., "Human Interferons And Intralesional Therapy Of Melanoma And Breast Carcinoma", Amer. Assoc. For Cancer Research, Abs nr. 993, p. 246 (1979)). Significantly, some cell lines which resist the anticellular effects of Le IF remain sensitive to F IF. This differential effect suggests that F IF may be usefully employed against certain classes of resistant tumor cells which appear under selective pressure in patients treated with high doses of

growth of many classes of tumors in many animals (The Interferon System, pp. 292-304). They, like other anti-tumor agents, seem most effective when directed against small tumors. The antitumor effects of animal IF are dependent on dosage and time but have been demonstrated at concentrations below toxic levels. Accordingly, numerous investigations and clinical trials have been and continue to be conducted into the antitumor and anticancer properties of HIFs. These include treatment of several malignant diseases such as osteosarcoma, acute myeloid leukemia, multiple myeloma and Hodgkin's disease (Texas Reports, pp. 429-35). In addition, F IF has recently been shown to cause local tumor regression when injected into subcutaneous tumoral nodules in melanoma and breast carcinoma-affected patients (T. Nemoto et al., Human Interferon And Intralesional Therapy of Melanoma And Breast Carcinoma", Amer. Assoc. For Cancer Research, Abs nr. 993, p. 246 (1979)).

Significantly, some cell lines which resist the anti-cellular effects of Le IF remain sensitive to F IF. This differential effect suggests that F IF may be usefully employed against certain classes of resistant tumor cells which appear under selective pressure in patients treated with high doses of Le IF (T. Kuwata et al., supra; A. A. Creasy et al., ~~"Regulatory Functions of Interferons" a~~ N. Y. Acad. Sci., ~~Abstract nr. 17~~ (1979)). Although the results of these clinical tests are encouraging, the antitumor and anticancer applications of HIF have been severely hampered by lack of an adequate supply of purified HIF.

"The Role of G. G. Assay In The Inhibition of Tumor Cell Growth By Interferon", Abstracts, Conference On Regulatory Functions Of Interferons, Oct. 23-25,

At the biochemical level IFs induce the formation of at least 3 proteins, a protein kinase (B. Lebleu et al., "Interferon, Double-Stranded RNA And Protein Phosphorylation", Proc. Natl. Acad. Sci. USA, 73, pp. 3107-11 (1976); A. G. Hovanessian and I. M. Kerr, "The (2'-5') Oligoadenylate (ppp A2'-5A2'-5'A) Synthetase And Protein Kinase(s) From Interferon-Treated Cells", Eur. J. Biochem., 93, pp. 515-26 (1979)), a (2'-5')oligo(A) polymerase (A. G. Hovanessian

comprises the steps of producing a single-stranded DNA copy (cDNA) of a purified messenger RNA (mRNA) template for the desired protein; converting the cDNA to double-stranded DNA; linking the DNA to an appropriate site in an appropriate cloning vehicle to form a recombinant DNA molecule and transforming an appropriate host with that recombinant DNA molecule. Such transformation may permit the host to produce the desired protein. These include, for example, Le IF (C. Weissmann et al., Seminar, Massachusetts Institute of Technology, January 16, 1980). In addition, recombinant DNA technology has been employed to produce a plasmid said to contain a gene sequence coding for F IF (T. Taniguchi et al., "Construction and Identification of a Bacterial Plasmid Containing the Human Fibroblast Interferon Gene Sequence", Proc. Japan Acad., 55, Ser. B, pp. 464-69 (1979).

However, in neither of the foregoing has the actual gene sequence of F IF been described and in neither has that sequence been compared to the initial amino acid sequence or amino acid composition of authentic F IF. The former Weissmann work is directed only to Le IF, which is distinct chemically, biologically and immunologically from Class I F IF (cf. supra). The latter report is Taniguchi et al. based solely on hybridization data. These latter data do not enable one to determine if the selected clone contains the complete or act gene sequence for F IF or if that the cloned gene sequence will be able to be expressed from F IF in bacteria. Hybridization only establishes that a particular DNA insert is to some extent homologous with and complementary to a mRNA component of the poly(A) RNA that induces interferon activity when injected into oocytes. Moreover, the extent of the homology is dependent on the hybridization conditions chosen for the screening process. Therefore, hybridization to a mRNA component of poly(A) RNA alone does not demonstrate that the selected DNA sequence is a sequence which codes for F IF or a polypeptide which displays the immunological or biological activity of F IF.

DISCLOSURE OF THE INVENTION

The present invention avoids the uncertainties referred to by providing at least one recombinant DNA molecule characterized by a structural gene whose nucleotide sequence is substantially consistent with the known amino acid composition and sequences of authentic F IF.

By virtue of this invention, it is therefore possible to obtain a structural gene that codes for a polypeptide whose amino acid sequence and composition is substantially consistent with authentic F IF. Replication of these genes in appropriate recombinant DNA molecule-host combinations permits the production of large quantities of these genes. These genes are useful, either as produced in the host or after appropriate derivatization or modification, in compositions and methods for detecting and improving the production of these products themselves and in selecting other gene related thereto.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 is a schematic outline of one embodiment of a process of this invention for preparing a mixture of recombinant DNA molecules, some of which are characterized by inserted DNA sequences that characterize this invention.

Figure 2 is a schematic outline of the initial clone screening process of this invention.

Figure 3 is a schematic outline of one embodiment of a clone screening process using DNA sequences prepared in accordance with the invention.

Figure 4 is a restriction map of one of the clones of the invention; the absolute position of each restriction site in this clone has not been determined.

DNA Sequence--A linear array of nucleotides connected one to the other by phosphodiester bonds between the 3' and 5' carbons of adjacent pentoses.

Codon--A DNA sequence of three nucleotides (a triplet) which encodes through mRNA an amino acid, a translation start signal or a translation termination signal. For example, the nucleotide triplets TTA, TTG, CTT, CTC, CTA and CTG encode for the amino acid leucine ("Leu"), TAG, TAA and TGA are translation stop signals and ATG is a translation start signal.

Reading Frame--The grouping of codons during translation of mRNA into amino acid sequences. During translation the proper reading frame must be maintained. For example, the sequence GCTGGTTGTAAG may be translated in three reading frames or phases, each of which affords a different amino acid sequence:

GCT GGT TGT AAG--Ala-Gly-Cys-Lys

G CTG GTT GTA AG--Leu-Val-Val

GC TGG TTG TAA G--Trp-Leu-(STOP)

Polypeptide--A linear array of amino acids connected one to the other by peptide bonds between the α-amino and carboxy groups of adjacent amino acids.

Genome--The entire DNA of a cell or a virus. It includes inter alia the structural genes coding for the polypeptides of the substance, as well as operator, promoter and ribosome binding and interaction sequences, including sequences such as the Shine-Dalgarno sequences.

Structural Gene--A DNA sequence which encodes through its template or messenger RNA ("mRNA") a sequence of amino acids characteristic of a specific polypeptide.

Transcription--The process of producing mRNA from a structural gene.

Translation--The process of producing a polypeptide from mRNA.

Expression--The process undergone by a structural gene to produce a polypeptide. It is a combination of transcription and translation.

Plasmid--A nonchromosomal double-stranded DNA sequence comprising an intact "replicon" such that the plasmid is replicated in a host cell. When the plasmid is placed within a unicellular organism, the characteristics of that organism may be changed or transformed as a result of the DNA of the plasmid. For example, a plasmid carrying the gene for tetracycline resistance (Tet^R) transforms a cell previously sensitive to tetracycline into one which is resistant to it. A cell transformed by a plasmid is called a "transformant".

Phage or Bacteriophage--Bacterial virus many of which consist of DNA sequences encapsidated in a protein envelope or coat ("capsid").

Cloning Vehicle--A plasmid, phage DNA or other DNA sequences which are able to replicate in a host cell, characterized by one or a small number of endonuclease recognition sites at which such DNA sequences may be cut in a determinable fashion without attendant loss of an essential biological function of the DNA, e.g., replication, production of coat proteins or loss of promoter or binding sites, and which contain a marker suitable for use in the identification of transformed cells, e.g., tetracycline resistance or ampicillin resistance. A cloning vehicle is often called a vector.

Cloning--The process of obtaining a population of organisms or DNA sequences derived from one such organism or sequence by asexual reproduction.

Recombinant DNA Molecule or Hybrid DNA--A molecule consisting of segments of DNA from different genomes which have been joined end-to-end outside of living cells and have the capacity to infect some host cell and be maintained therein.

Expression Control Sequence--A sequence of nucleotides that controls and regulates expression of structural genes when operatively linked to those genes.

Referring now to Figure 1, we have shown therein a schematic outline of one embodiment of a process for preparing a mixture of recombinant DNA molecules, some of which include inserted DNA sequences that characterize this invention.

PREPARATION OF POLY(A) RNA CONTAINING HUMAN FIBROBLAST INTERFERON mRNA (F IF mRNA)

The RNA used in this invention was extracted from human VGS cells, a diploid fibroblast cell line which can be propagated in monolayer cultures at 37°C. ~~Interferon~~ F IF is produced in these cells on induction with poly(I,C) ~~and~~ in the presence of cycloheximide.

For a typical RNA isolation, each of 20 roller bottles of diploid VGS cells in confluent monolayer were "primed" overnight with 100 units/ml F IF and the cultures induced for 1 h with 100 µg/ml poly(I,C) and 50 µg/ml cycloheximide, incubated with cycloheximide (50 µg/ml) for 4 h, harvested by scraping into phosphate-buffered saline and spun down. The cells were lysed for 15 min at 0°C ~~and~~ to remove the intact nuclei containing the DNA and to isolate the cytoplasmic RNA by suspending them in hypotonic buffer (10 M Tris-Cl (pH 7.4), 10 mM NaCl and 1.5 mM $MgCl_2$) and adding NP40 to 1%. Nuclei were removed by pelleting in a Sorvall SS-34 rotor for 5 min at 3000 rpm. Sodium dodecyl sulphate ("SDS") and EDTA were added to the supernatant to 1% and 10 mM, respectively, and the mixture extracted 5 times with 2x vol of 1:1 redistilled phenol and chloroform-isoamylic alcohol (25:1), the aqueous phases containing the RNA being separated by centrifugation in a Sorvall SS-34 rotor at 8000 rpm for 10 min after each extraction. The RNA was precipitated from the aqueous phase by addition of 1/10 vol 2 M sodium acetate (pH 5.1) and 2.5 vol of ethanol. Usually, 60 to 90 µg of total cytoplasmic RNA was obtained per roller bottle.

Other procedures to extract the cytoplasmic RNA have also been used. For example, the cells were totally lysed after homogenization in 0.2 M Tris-HCl (pH 9.0) 50 mM NaCl, 20 mM EDTA and 0.5% ~~sodium dodecyl sulphate~~ SDS and extracted with phenol-chloroform as above or the washed cells were suspended in 400 ul 0.1 M NaCl, 0.01 M Tris-HCl (pH 7.5), and 0.001 M EDTA ("NTE buffer") and 2.5 ml 4 M guanidinium-isothiocyanate and 1 M β-mercaptoethanol in 20 mM sodium acetate (pH 5.0) were added and the cells homogenized. The lysate was layered on a 1.3-ml 5-7 M CsCl cushion in a Beckman SW-60 Ti nitrocellulose tube, spun for 17 h at 39000 rpm to pellet the RNA and separate it from DNA, proteins and lipids and the RNA extracted once with phenol-chloroform (Reynolds et al., "Interferon Activity Produced By Translation Of Human Interferon Messenger RNA In Cell-Free Ribosomal Systems And In Xenopus Oocytes", Proc. Natl. Acad. Sci. USA, 72, pp. 4881-885 (1975); T. Moser et al., "Characterization Of Interferon Messenger RNA From Human Lymphoblastoid Cells", J. Gen. Virol., 44, pp. 231-34 (1979)).

The total RNA was assayed for F IF mRNA by injection into the cytoplasm of Xenopus laevis oocytes and determining the ~~interferon~~ F IF activity induced therein (Reynolds et al., supra). The assay was conducted by dissolving the RNA in water and injecting about 50 ul into each oocyte. The oocytes were incubated overnight at room temperature in Barth medium (J. Gurdon, J. Embryol. Exper. Morphol., 20, pp. 401-14 (1968)), homogenized in part of the medium, the debris removed by centrifugation, and the F IF activity of the supernatant determined. Detection of F IF activity was by reduction of virus-induced cytopathic effect (W. E. Stewart and S. E. Sulkin, Proc. Soc. Exp. Biol. Med., 123, pp. 650-53 (1966)). The challenge virus was vesicular stomatitis virus (Indiana

strain) and the cells were human diploid fibroblasts trisomic for chromosome 21 to afford higher F IF sensitivity. F IF activity is expressed relative to the IF reference standard 69/19.

Poly(A) RNA containing F IF mRNA was isolated from the cytoplasmic RNA by adsorption to oligo(dT)-cellulose (type 7; P-L Biochemicals) in 0.4 M NaCl, 10 mM Tris-HCl (pH 7.8), 10 mM EDTA and 0.2% ~~sodium dodecyl sulphate~~ SDS for 10 min at room temperature. RNA aggregation was minimized by heating the RNA for 2 min at 70°C prior to adsorption. After washing the cellulose with the above-mentioned buffer, the poly(A) RNA fraction was eluted with 10 mM Tris-HCl (pH 7.8), 1 mM EDTA and 0.2% ~~sodium dodecyl sulphate~~ SDS. It usually comprised 4-5% of the total RNA, as measured by optical density at 260 nm.

A further purification to enrich the poly(A) RNA in F IF mRNA was effected by formamide-sucrose gradients (T. Pawson et al., "The Size of Rous Sarcoma Virus mRNAs Active in Cell-Free Translation", Nature, 268, pp. 416-20 (1977)). These gradients gave much higher resolution than the nondenaturing sucrose gradients. Usually about 80 µg poly(A) RNA was dissolved in 50% formamide, 100 mM LiCl, 5 mM EDTA, 0.2% ~~sodium dodecyl sulphate~~ SDS and 10 mM Tris-HCl (pH 7.4), heated at 37°C for 2 min to prevent aggregation and loaded on a 5-20% sucrose gradient in a Beckman SW-60 Ti polyallomer tube. After centrifugation at 20°C for 4 1/2 h at 60000 rpm in the Beckman SW-60 Ti rotor with total ^{14}C-labeled eukaryotic RNA serving as size markers, the gradient was fractionated and the optical density of the fractions ~~was~~ determined. All RNA fractions were precipitated twice with 0.5 M NaCl and 2.5 vol ethanol and assayed for interferon mRNA activity as described above.

Alternatively, the oligo(dT)-adsorbed mRNA (60 µg) was fractionated by electrophoresis in a 4% polyacrylamide gel in 7 M urea, 0.1% ~~sodium dodecyl sulphate~~ SDS,

CLONING OF DOUBLE-STRANDED DNA

A wide variety of host/cloning vehicle combinations may be employed in cloning the double-stranded cDNA prepared in accordance with this invention. For example, useful cloning vehicles may consist of segments of chromosomal, non-chromosomal and synthetic DNA sequences, such as various known derivatives of SV40 and known bacterial plasmids, e.g., plasmids from E. coli including col E1, pCR1, pBR322, pMB9 and their derivatives, wider host range plasmids, e.g., RP4, phage DNA, e.g., the numerous derivatives of phage λ e.g., NM 989, and other DNA phages, e.g., M13 and fd, and vectors derived from combinations of plasmids and phage DNAs such as plasmids which have been modified to employ phage DNA or other expression control sequences or yeast plasmids such as the 2 u plasmid or derivatives thereof. Useful hosts may include bacterial hosts such as E. coli HB 101, E. coli X1776, E. coli X2282, E. coli MRC1 and strains of Pseudomonas, Bacillus subtilis, Bacillus stearothermophilus and other bacilli, yeasts and other fungi, animal or plant hosts such as animal (including human) or plant cells in culture or other hosts. Of course, not all host/vector combinations may be equally efficient. The particular selection of host/cloning vehicle combination may be made by those of skill in the art after due consideration of the principles set forth without departing from the scope of this invention.

Furthermore, within each specific cloning vehicle, various sites may be selected for insertion of the double-stranded DNA. These sites are usually designated by the restriction endonuclease which cuts them. For example, in pBR322 the PstI site is located in the gene for β-lactamase, between the nucleotide triplets that code for amino acids 181 and 182 of that protein. This site was employed by C. Weissmann et al., supra, in their synthesis of polypeptides displaying an immunological or biological activity of LIF. One of the two HindII endonuclease recognition sites is between the triplets coding for amino acids 101 and 102 and one of the several Taq sites at the triplet coding for amino acid 45 of β-lactamase in pBR322. In similar fashion, the EcoRI site and the PvuII site in this plasmid lie outside of any coding region, the EcoRI site being located between the genes coding for resistance to tetracycline and ampicillin, respectively. This site was employed by T. Taniguchi et al., supra, in

their recombinant synthetic scheme. These sites are well recognized by those of skill in the art. It is, of course, to be understood that a cloning vehicle useful in this invention need not have a restriction endonuclease site for insertion of the chosen DNA fragment. Instead, the vehicle could be joined to the fragment by alternative means.

The vector or cloning vehicle and in particular the site chosen therein for attachment of a selected DNA fragment to form a recombinant DNA molecule is determined by a variety of factors, e.g., number of sites susceptible to a particular restriction enzyme, size of the protein to be expressed, susceptibility of the desired protein to proteolytic degradation by host cell enzymes, contamination of the protein to be expressed by host cell proteins difficult to remove during purification, expression characteristics, such as the location of start and stop codons relative to the vector sequences, and other factors recognized by those of skill in the art. The choice of a vector and an insertion site for a particular gene is determined by a balance of these factors, not all selections being equally effective for a given case.

Although several methods are known in the art for inserting foreign DNA into a cloning vehicle or vector to form a recombinant DNA molecule, the method preferred in accordance with this invention is characterized by digesting the plasmid (in particular pBR322) with that restriction enzyme specific to the site chosen for the insertion (in particular PstI) and adding dA tails to the 3' termini by terminal transferase. In similar fashion, the double-stranded cDNA is elongated by the addition of dT tails to the 5' termini to allow joining to the tailed plasmid. The tailed plasmid and cDNA are then annealed to insert the cDNA in the appropriate site of the plasmid and to circularize the hybrid DNA, the complementary character of the tails permitting their cohesion (Figure 1). The resulting recombinant DNA molecule now carries a gene at the chosen restriction site (Figure 1). This method of dA-dT tailing for insertion is described by D.A. Jackson et al., "Biochemical Methods for Inserting New Genetic Information into DNA of Simian Virus 40 : Circular SV40 DNA Molecules Containing Lambda Phage Genes and the Galactose Operon of Escherichia coli", Proc. Natl. Acad. Sci. USA, 69, pp. 2904-2909 (1972) and R. Devos et al., supra. It results in about 3

prepared. Again, only a very few of these clones will contain the gene for FIF or fragments thereof (Figure 1). The preferred host in accordance with this invention is E. coli BH 101.

1. Preparation of PstI-Cleaved, dCMP-elongated pBR322

Plasmid pBR322 (30 µg) was digested with 15 units PstI endonuclease (New England Biolabs) in 150 µl 10 mM Tris-HCl (pH 7.6), 7 mM $MgCl_2$, 50 mM NaCl, 7 mM 2-mercaptoethanol, 300 µg/ml of bovine serum albumin ("BSA") (Calbiochem). After 2 h at 37°C, the mixture was extracted several times with 1 vol phenol and 2-3 vol ether and precipitated with 2.5 vol ethanol; 0.2 M Sodium Acetate solution

Addition of homopolymeric dA tails (Figure 1) by terminal deoxynucleotidyl transferase (TdT) (purified according to L. Chang and F.J. Bollum, "Deoxynucleotide-polymerizing Enzymes of Calf Thymus Gland", J. Biol. Chem., 246, pp. 909-16 (1971)) was done in a 50-µl reaction volume containing 0.14 M potassium cacodylate, 30 mM Tris-HCl (pH 6.8), 1 mM $CoSO_4$, 0.2 µg/µl heat-inactivated bovine serum albumin, 0.8 mM DTT, 0.2 mM dATP and some α-^{32}P-dATP. Incubation was at 37°C for 5 min before EDTA was added to 10 mM and SDS to 0.1 % and the mixture extracted with phenol and chromatographed on Sephadex G50 in TE buffer. The void fractions, containing the linearized and elongated pBR322, were further purified by adsorption in 10 mM Tris-HCl (pH 7.8), 1 mM EDTA and 0.4 M NaCl to oligo(dT)cellulose. After extensive washing, the desired fractions were eluted with 10 mM Tris-HCl (pH 7.8) and 1 mM EDTA.

2. Preparation of dT-elongated DNA

Double-stranded DNA was elongated with dTMP residues in similar fashion to that described above for dA tailing of pBR322, except that dTTP and some ^{3}H-dTTP replaced the dATP and α-^{32}P-ATP. Purification on oligo(dT) cellulose was, of course, omitted. As before, the dT-elongated DNA is a mixture of different species, only a very few of which are IF-related (Figure 1).

3. Preparation of Ca^{++}-Treated E.coli HB101

Ca^{++}-treated E.coli HB101 was prepared by the method of

E. M. Lederberg and S.N. Cohen, "Transformation of Salmonella Typhimurium By Plasmid Deoxyribonucleic Acid", J. Bacteriol., 119, pp. 1072-74 (1974) by inoculating the E.coli HB101 (a gift from H. Boyer) into 5 ml LB medium (10 parts bacto-tryptone, 5 parts yeast extract and 5 parts NaCl per liter) and cultures grown overnight at 37°C. The fresh cultures were diluted 1/100 in 20 ml LB medium and grown to a density of about 2 x 10^8 bacteria per ml, quickly chilled in ice and pelleted at 6000 rpm for 5 min in a Sorvall SS34 rotor at 4°C. The cells, kept at 0-4°C, were washed with 20 ml 100 mM $MgCl_2$, repelleted by centrifugation and suspended in 10 ml 100 mM $CaCl_2$. After 20 min in ice, the cells were repelleted and resuspended in 2 ml 100 mM $CaCl_2$ and maintained at 0°C for 15 min. Aliquots (200 µl), supplemented with glycerol to 11%, could be stored for several months at -80°C without loss of activity (D.A. Morrison, "Transformation in Escherichia coli : Cryogenic Preservation of Competent Cells", J. Bacteriol., 132, pp. 349-51 (1977)).

4. Annealing of dA-elongated pBR322 and dT-elongated DNA

The vector's and DNA insert's complementary dA- and dT-tails permit annealing to form the desired hybrid plasmid or recombinant DNA molecule. For this purpose, the dA-tailed PstI-cleaved pBR322 vector and the mixture of sized dT-tailed cDNAs were dissolved in TSE buffer (10 mM Tris-HCl (pH 7.6), 1 mM EDTA, 100 mM NaCl) to 1.5 µg/ml plasmid and to a molar ratio of plasmid to DNA insert of 1.5 to 2.0. After heating to 65°C for 10 min, the mixture was cooled slowly to room temperature over 4 h.

The product is, of course, a large mixture of different recombinant DNA molecules and some cloning vehicles without inserted DNA sequences. However, each recombinant DNA molecule contains a cDNA segment at the PstI site. Each such cDNA segment may comprise a gene or a fragment thereof. Only a very few of the cDNA segments code for FIF or a portion thereof (Figure 1). The vast majority code for one of the other proteins or portions thereof whose mRNA's were part of the poly(A) RNA used in the process of this invention (Figure 1).

5. Transfection Of E.coli HB101 With The Annealed Hybrid Plasmids

P3 containment facilities were used for the transfection process and all subsequent steps in which the resulting transformed bacteria were handled. Aliquots (90 µl or less) of the above mixture were cooled to 0°C and 1 M $CaCl_2$ added to 0.1 M. Aliquots (100 µl or less) of this solution were added to 200 µl Ca^{++}-treated E. coli HB101 in ice and after standing at 0°C for 30 min, the cells were heat-shocked for 5 min at 37°C and cooled again at 0°C for 15 min. After addition of 2 ml LB-medium, the cells were incubated at 37°C in a shaking water bath for 30 to 45 min and the bacterial suspension plated out onto 1.2 % agar plates, containing LB medium supplemented with 10 µg/ml tetracycline.

Since plasmid pBR322 includes the gene for tetracycline resistance, E.coli hosts which have been transformed with a plasmid having that gene intact will grow in cultures containing that antibiotic to the exclusion of those bacteria not so transformed. Therefore, growth in tetracycline-containing culture permits selection of hosts transformed with a recombinant DNA molecule or recyclized vector.

After 24 h at 37°C, individual colonies were picked and suspended in 100 µl LB medium (supplemented as above) in the wells of microtiter plates (Dynatech). After incubation at 37°C overnight, 11 µl dimethylsulfoxide were mixed into each well and the trays sealed with adhesive tape. The plates were stored at -20°C and a library of 17,000 individual clones of transformed E. coli HB101 was prepared. This library was derived from 270 fmoles (128 ng) dT-tailed cDNA inserts, which in turn were synthesized from 4.4 µg gradient purified poly(A) RNA. About 98 % of the clones of this library (band on representative fractions) were sensitive to carbanicillin (a more stable ampicillin-derivative). Therefore, about 98% of the library contained a plasmid having an insert in the PstI-site of the ß-lactamase gene of pBR322 only about 2% contained a recyclized vector without insert.

These 17,000 clones contain a variety of recombinant DNA molecules representing complete or partial copies of the mixture of mRNAs in the poly(A) RNA preparation from FIF-producing cells (Figure 2). The majority of these will contain only a single recombinant DNA molecule. Only a very few of these recombinant

DNA molecules are related to FIF. Accordingly, the clones must be screened to separate the FIF-related clones from the others.

SCREENING FOR A CLONE CONTAINING F IFcDNA

There are several approaches to screen for bacterial clones containing ~~interferon~~ F IF cDNA. ~~(IFcDNA)~~. These include, for example, RNA selection hybridization (Alwine et al., infra), differential hybridization (T.P. St. John and R.W. Davis, "Isolation of Galactose-Inducible DNA Sequences from Saccharomyces Cerevisiae by Differential Plaque Filter Hybridization", Cell, 16, pp. 443-452 (1979)); ~~Hoeymackers et al., infra~~, hybridization with a synthetic probe (B. Noyes et al., "Detection And Partial Sequence Analysis Of Gastrin mRNA By Using An Oligodeoxynucleotide Probe", Proc. Natl. Acad. Sci. USA, 76, pp. 1770-1774 (1979)) or screening for clones that produce the desired protein by immunological (A.C.Y. Chang et al., supra) assays. We have chosen RNA selection hybridization as being the most convenient and promising method for primary screening.

(L. Villa-Komaroff et al., "A Bacterial Clone Synthesizing Proinsulin", Proc. Natl. Ac Sci USA, 75, pp 3727-31 (1978)) or Analogue

"Phenotypic Expression In E. coli Of A DNA Sequence Coding For Mouse Dihydrofolate Reductase", Nature, 285, pp 497-24 (1978))

A. RNA Selection Hybridization Assay

1. Overview Of The Initial Assay

Referring now to Figure 2, the recombinant DNA molecules were isolated from a culture of a mixture of about 46 clones sensitive to carbenicillin and resistant to tetracycline from the above library of clones (two mixtures of 2 clones shown in Figure 2) (Step A). The recombinant DNA molecules were cleaved, ~~denatured~~ and hybridized to total RNA containing F IFmRNA prepared as before (Step B). All recombinant DNA molecule-total RNA hybrids were separated from the non-hybridized total RNA (Step C). The hybridized total RNA was recovered from the hybrids and purified (Step D). The recovered RNA was assayed for F IFmRNA activity as above (Step E). If, and only if, the mixture of recombinant DNA molecules contains a recombinant DNA molecule having an inserted nucleotide sequence capable of hybridizing to the F IFmRNA in the total RNA, under stringent hybridization conditions, will the mRNA released from that hybrid cause the formation of F IF in oocytes, because mRNA released from any other recombinant DNA molecule-total RNA hybrid will not be F IF-related. If a group of 46 clones gave a positive response, the clones were regrouped ~~in~~ 4 ~~lots~~ subgroups of 8 and 2 ~~lots~~ subgroups of 7), and into 6 subgroups (

each subgroup assayed as before. This process was continued until a single clone responding to this assay was identified.

There is no assurance that the recombinant DNA molecules and bacterial clones transformed therewith, which are thus identified, contain the complete F IFcDNA sequence of F IF or even that the DNA sequence actually codes for F IF. However, the recombinant DNA molecules will certainly contain extensive nucleotide sequences complementary to the F IFmRNA coding sequence. Therefore, the recombinant DNA molecule may at least be used as a source of a probe to screen rapidly other recombinant DNA molecules and clones transformed with them to identify further sets of clones which may contain nucleotide an authentic and complete F IF nucleotide coding sequence. The sequencing of these by insert DNA correlated to fragment of these hybrid plasmids may also be determined and compared the compositional amino acid composition and initial sequence reported for authentic F IF (sig

2. Execution Of The Initial Assay

Step A - Preparation Of The Recombinant DNA Molecule Mixture

Replicas of a microtiter plate containing 96 clones from the above library of clones were made on LB-agar plates, one containing 10 ug/ml tetracycline and the other supplemented with ~96 100 ug/ml carbenicillin. In this manner, two sets of about 45 clones, resistant to tetracycline and sensitive to carbenicillin, were picked and grown separately overnight at 37°C in 100 ml LB medium, containing 10 ug/ml tetracycline. These cultures were pooled, spun down in a Sorvall GS-3 rotor at 8000 rpm for 10 min, washed twice with TES buffer (50 mM Tris-Hcl (pH 8), 5 mM EDTA, 5 mM NaCl) and resuspended in 40 ml TES per l of initial culture volume. The cells were lysed with lysozyme-Triton X-100 (M. Kahn et al., "Plasmid Cloning Vehicles Derived From Plasmids ColE1, F, R6K And RK2", in Methods of Enzymology, 68, : Recombinant DNA (R. Wu, ed.) (1980) in press). Forty ml of the TES suspended cells were combined with 20 ml 10% sucrose in 50 mM Tris-HCl (pH 8) and lysozyme to 1.3 mg/ml and allowed to stand at room temperature for 20 min. To this suspension were added 1 ml 0.5 M EDTA-NaOH (pH 8), 8 ml 0.2% Triton X-100, 25 mM EDTA, 50 mM Tris-HCl (pH 8) and the lysis completed at room temperature for 30 min. Cellular debris and most of the chromosomal DNA were removed by pelleting in a Beckmann SW27 rotor at 24000 rpm for 45 min. The supernatant was

cooled in ice, combined with 1/3 vol 40 % polyethylene glycol 6000 - 2 M NaCl and allowed to stand overnight at 0°C. The resulting precipitate was collected in a Sorvall HB4 rotor at 5000 rpm for 10 min at 4°C and dissolved in TES buffer. The solution, with 0.2 vol 10 mg/ml ethidium bromide (Serva) and CsCl to 1 g/ml, was centrifuged in a Beckmann R60 Ti-rotor at 40000 rpm for at least 48 h, one polyallomer tube usually being sufficient for the lysate from 1-2 l of original culture volume. Two DNA bands could be visualized in the tube by UV-illumination. The band of highest density corresponds to plasmid form I DNA, the second band corresponds to form II and form III plasmid DNAs and some chromosomal DNA. The first band was cut from the tube, ethidium bromide removed by six isoamyl alcohol extractions, and the aqueous phase diluted with 3 vol water-supplemented with up to 0.2 M sodium acetate (pH 5.1) before DNA precipitation with 2.5 vol ethanol. The DNA was redissolved, extracted with phenol and again precipitated with ethanol. The quality of the DNA was monitored by electrophoresis on a 1% agarose gel in 40 mM Tris-OAc (pH 7.8), 20 mM sodium acetate, 2 mM EDTA, followed by (ethidium bromide staining). If the DNA contained was contaminated with form II or form III DNAs, it was further purified by neutral sucrose-gradient centrifugation: 300 µg DNA in 10 mM Tris-HCl (pH 7.6) and 1 mM EDTA was loaded on a 36-ml 5-20% sucrose gradient in 10 mM Tris-HCl (pH 7.6), 1 mM EDTA, 1 M NaCl, centrifuged in polyallomer tubes for 16 h at 24000 rpm in a Beckmann SW27-rotor at 18°C and the DNA containing fractions (OD_{260}) pooled and precipitated with sodium acetate-ethanol.

Step B - Hybridization Of The DNA With Total RNA

About 150 µg DNA, thus prepared, was combined with some uniformly labelled ^{32}P-marker DNA and 2 µg pSTNV-1 DNA (a recombinant plasmid containing a full size cDNA copy of satellite tobacco necrosis virus ("STNV")-RNA; J. Van Emmelo et al., "Construction And Characterization Of A Plasmid Containing A Nearly Full-Size DNA Copy Of Satellite Tobacco Necrosis Virus RNA", J. Mol. Biol., submitted for publication) as internal control, sheared by sonication in an MSE sonicator and precipitated with sodium acetate-ethanol.

A diazobenzyloxymethyl (DBM)-cellulose solid matrix (cf., J.C. Alwine et al., "Method For Detection Of Specific RNAs In Agarose Gels By Transfer To Diazobenzyl Oxymethyl Paper and Hybridization With DNA Probes", Proc. Natl. Acad. Sci. USA, 74,

pp. 5350-54 (1977)) was prepared according to the method of J.C. Alwine et al., "Detection of Specific RNAs or Specific Fragments of DNA by Gel Transfer and Hybridization", Method in Enzymology, 68: Recombinant DNA (R. Wu, ed.) (1980) (in press). For a paper matrix, a sheet of Whatman 540 paper was evenly soaked in 2-3 ml 1-(m-nitrobenzyloxy)methyl pyridinium chloride (NBPC/BDH) - 0.7 ml sodium acetate trihydrate - 2.8 ul water per cm^2, incubated at 60°C until dry and for a further 10 min, and baked at 130-135°C for 30-40 min. After washing several times with water (about 20 min), 3 times with acetone (about 20 min) and drying, the paper was incubated at 60°C for 30 min in 0.4 ml 20% sodium dithionite-water with occasional shaking. The paper was again washed four times with water, once with 30% acetic acid for 5 min and four times with water, transferred for 30 min at 0°C to 0.3 ml per cm^2 ice-cold 1.2 M HCl to which 10 mg/ml fresh $NaNO_2$ had been added immediately before use, and washed twice quickly with ice-cold water and once with 80% dimethyl sulfoxide (spectrophotometric grade, Merck) - 20% 25 mM sodium phosphate (pH 6.0). For a powder matrix essentially the same procedure was followed using micro granular cellulose powder (Whatman CC31), the quantities being expressed against the corresponding weight of the cellulose matrix.

Initially, we used a powder matrix because the capacity for binding was higher, so relatively smaller volumes for hybridization, washes and elution could be used. Subsequently, we used a paper matrix for individual clone screening. Use of paper permits efficient elution with water which proved superior for the later assay of F IPmRNA. (for powder) (for paper)

The DNA (50-100 µg, ~~for paper~~ 3-4 µg) prepared above was dissolved in 25 mM sodium phosphate (pH 6.0) heated for 1 min, chilled and four vol DMSO added. Coupling to the matrix (50 mg (powder) or a disc (20 mm dia.)) usually proceeded over a weekend at 4°C with continuous mixing. The volume of the DNA was kept rather small to allow close contact with the matrix and thereby enhance efficient coupling of the DNA to the matrix. After coupling, the matrix was washed four times with water, four times with 0.4 N NaOH at 37°C for 10 min each, again four times with water at room temperature and finally twice with hybridization buffer (50% formamide (deionized, Baker), 40 mM piperazine-N,N'-bis(2-ethane sulfonic acid) (pH 6.4) ("PIPES", Sigma), 1 mM EDTA, 0.6 M NaCl and 0.1 % SDS) at 4°C. Coupling efficiencies were measured by ^{32}P-radioactivity.

Twenty µg total RNA, prepared as before, and 50 ng STNV-RNA were dissolved in 250 µl (50 µl for paper matrix) hybridization buffer and added to the DNA coupled matrix. The matrix was heated to 70°C for 2 min and held at 37°C overnight with gentle mixing.

Step C - Separation Of Hybridized Total RNA-DNA From Non-Hybridized Total RNA

After centrifugation of ~~the~~ ~~above~~ powder matrix, the unhybridized RNAs were removed and the matrix washed seven times with (0.28 ml) 50% formamide, 10 mM PIPES (pH 6.4), 1 mM EDTA, 0.3 M NaCl and 0.1% SDS, the lower salt content of these washes destabilizing non-specific RNA-DNA binding. Each wash was followed by centrifugation and resuspension of the matrix in the buffer. For subsequent assay, the first wash was pooled with the unhybridized DNA ("Fraction 1") and washes 2-4 ("Fraction 2") ~~as~~ and 5-7 ("Fraction 3") were pooled. In ~~those~~ hybridization to a paper matrix, a similar procedure was ~~employed~~ except that (water) was used for each wash. ~~(0.14 ml)~~

Step D - Purification Of Hybridized Total RNA

The hybridized total RNA-DNA was eluted from ~~the~~ powder matrix with 900 µl 99% formamide, 0.2% SDS at 70°C for 2 min and chilled in ice (A.G. Smith, personal communication). The hybrized total RNA-DNA was eluted from ~~the~~ paper matrix by (100 µl of ice cold water and) two 150 µl water elutions at 80°C for 2 min. For subsequent assay these elutions and the 100 µl wash were pooled ("Fraction 4").

To one-half of each fraction, 0.1 µg calf liver tRNA or ribosomal RNA were added (Fractions 1A, 2A, 3A and 4A) and to the other half 8 µg eukaryotic poly(A) RNA or ribosomal RNA were added (Fractions 1B, 2B, 3B, 4B). The fractions were (precipitated) by the addition of 0.5 M NaCl and 2.5 vol ethanol to remove traces of formamide and other impurities.

Step E - Determination Of F IFmRNA Activity

Fractions 1A, 2A, 3A and 4A were translated in (25 µl) nuclease-

(A)

After incubation, 25

Twenty-five µl [illegible] lysate, from above, were combined

with 1 µl 10% deoxycholate - 10% Triton X100 and 2 µl anti-serum - PBS

(1:9) and were heated at 37°C for 1 h. Twenty µl Staphylococcus

aureus Cowan I (freshly washed, S. W. Kessler et al., Rapid Isolation of

Antigens From Cells WITH A Staphylococcal Protein A-Antibody Adsorbent : Parameters

) in 10% 100mM NaCl, 10mM Tris-HCl (pH 7.4), 1mM

EDTA, 0.05% NP40 were added and the mixture maintained at 20°C

for 30 min and centrifuged in an Eppendorf 5412 centrifuge for 2 min.

The pellet was washed and centrifuged twice with PBS and the

final pellet dissolved in sample buffer and electrophoresed in 13%

polyacrylamide [illegible] described by U.K. Laemmli et al., "Cleavage of

Structural Proteins During The Assembly OF THE Head OF Bacteriophage

T4"

J. of Immunology, 111, pp 1639 - 1644 (1973)

treated rabbit reticulocyte lysate (R.B. Pelham and R.J. Jackson, "An Efficient mRNA-Dependent Translation System For Reticulocyte Lysates", Eur. J. Biochem., 7, pp. 247-56 (1976)) in the presence of ^{35}S-methionine, ~~immunoprecipitated with antiserum, electrophoresed in 13% polyacrylamide gel (W.K. Laemmli, Nature, 227, pp. 680-685 (1970)~~, and autoradiographed. Comparison of the STNV-RNA translation products in Fractions 1A and 4A provide an indication of the efficiency of hybridization and RNA degradation in the process. Fractions 1B, 2B, 3B and 4B were dissolved in 2 µl water and assayed in oocytes for F IFmRNA content as described above.

3. Subsequent Assay - Hybridization To Nitrocellulose Sheets

Some subsequent assays of individual clones were done on nitrocellulose sheets (M. Crochet et al., "Cloning Of An Almost Full-Length Chicken Conalbumin Double-Stranded cDNA", Nucleic Acids Res., 6, pp. 2435-2452 (1979)). The DNA was dissolved in 2M NaCl and 0.2 M NaOH, heated at 100°C for 1 min, chilled, and spotted on detergent free Millipore filters (pore size 0.45 um; 1 mm dia.). The filters were baked for 2 h at 80°C, washed in 0.3 M NaCl, 2 mM EDTA, 0.2% SDS, 10 mM Tris-HCl (pH 7.5) and dried at room tperature. The RNA was hybridized for 3 h at 47°C in 30% formamide, 0.5 M NaCl, 0.4% SDS, 2 mM EDTA, 50 mM PIPES (pH 7.5). Hybridization was stopped by dilution with 10 vol 0.1 M NaCl and the filters were washed several times in 15 ml of 0.3 M NaCl, 0.1% SDS, 2 mM EDTA, 10 mM Tris-HCl (pH 7.5) by shaking at 45°C and several times in the same solution without SDS at 4°C. Elution of the hybridized RNA-DNA was effected in 30 µl 5 mM potassium chloride at 100°C for 1 min.

4. Results Of The RNA Selection Hybridization Assay

Sixteen groups of about 46 clones were screened (Groups A-P). In six of the groups, Fraction 1B contained the only F IFmRNA activity, in eight of the groups no F IFmRNA was detected and in two groups (Groups C and P) F IFmRNA was observed in Fraction 4B. The positive assays are reported in the following format : logarithm of F IF units (calibrated against reference standard 69/19),

assay of Fraction 1B (background) and assay of Fraction 4B.

Group	Fraction 1B	Fraction 4B
C	1.0	0
	0.5	0.5
	0	0.2
O	0	0
	0.2	0.5

Group O was subdivided into 6 subgroups (O_1 to O_6; four of eight clones and two of seven clones) and hybridized and assayed as before, except a 400 ml culture per clone was used. The subgroups gave the following results, presented in the same format as above:

Subgroup	Fraction 1B	Fraction 4B
O_1	0	1.2
	0	1.5
	0	0.5
	0	0.5
	0.2	0.5
	0	1.2*
O_2	0.7	0
O_3	0.7	0
	0.5	0
O_4	0	0
O_5	0.5	0
O_6	0	0

* DPM paper method

Subgroup O_1 was subdivided into its individual clones (designated clones 1-8) and hybridized and assayed as before, except a 700 ml culture per clone was used :

Clone	Fraction 1B	Fraction 4B
$O_{1/1}$	0.2	0
	0.7	0
	0.7	0*
	1.0	0**

	Fraction 1B	Fraction 4B
$O_{1/2}$	1.2	0
	0.2	0*
	0.7	0**
$O_{1/3}$	1.2	0
	1.0	0.2*
	1.2	1.0(?)*
	1.2	0**
$O_{1/4}$	1.2	0
	1.2	0
	1.0	0*
	1.2	0**
$O_{1/5}$	0.7	0
	0.7	≤ 0.2*
	1.0	0**
$O_{1/6}$	0.7	0
	1.0	≤ 0.2*
	0.5	0**
$O_{1/7}$	0.5	0
	1.2	0*
	< 0.2	0.5**
$O_{1/8}$	0	1.7*
	< 0.2	1.2*
	0	0.7**
	0	1.0**

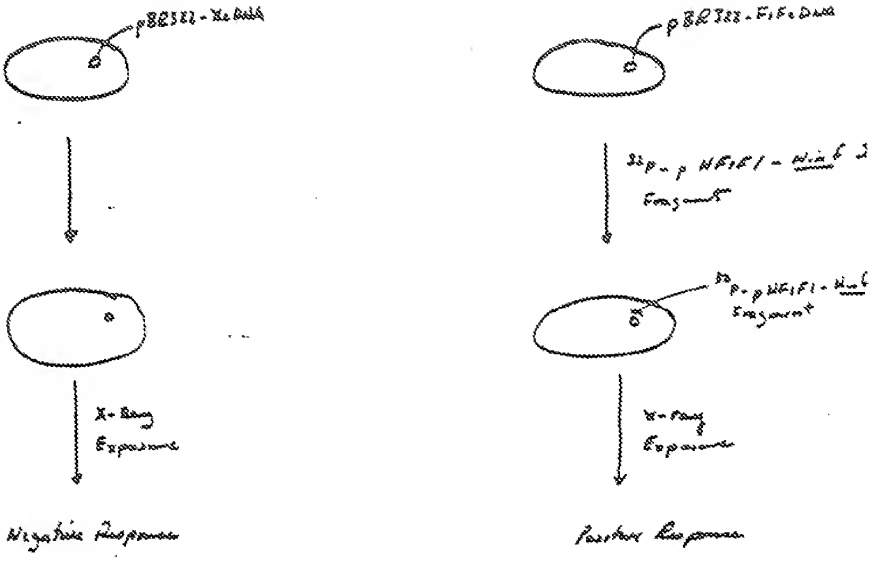
---

* DPM paper method

** Nitrocellulose sheets

Therefore, clone $O_{1/8}$ contains a recombinant DNA molecule capable of hybridizing F IFmRNA from total RNA containing F IFmRNA. Non-specific RNA-DNA binding is unlikely, because a comparison of fractions 1A and 4A revealed substantially no non-specific binding of STUV DNA.

FIG. 3



Filed on Behalf of Junior Party Sugano et al.
By: Nels T. Lippert (Reg. No. 25,888)
Jane M. Love, Ph.D. (Reg. No. 42,812)
William W. Kim, Ph.D. (Reg. No. 53,127)
WILMER CUTLER PICKERING
HALE AND DORR, LLP
399 Park Avenue
New York, NY 10022
Tel: 212.937.7233
Fax: 212.230.8888

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCE
(Administrative Patent Judge Sally Gardner Lane)

HARUO **SUGANO**,
MASAMI MURAMATSU, and TADATSUGU TANIGUCHI

Junior Party
(Patent 5,514,567 and 5,326,859; Application 08/463,757)
v.

DAVID V. **GOEDDEL**,
and ROBERTO CREA

Senior Party
(Application 07/374,311; Patent 5,460,811).

Patent Interference Nos. 105,334 and 105,337(SGL)
(Technology Center 1600)

DECLARATION OF CHARLES WEISSMANN, M.D., PH.D.

I, CHARLES WEISSMANN declare and state as follows:

1. I reside at 100 Sunrise Ave., Apt. 603E, Palm Beach, FL 33480.

2. I hold the degrees of M.D. and Ph.D. in Organic Chemistry from Zürich University, Zürich, Switzerland.

3. I presently hold the position of head of the Department of Infectology at The Scripps Research Institute in Palm Beach County, Florida.

4. I am a member of the Royal Society (United Kingdom) and National Academy of Science (USA).

5. My complete academic background, professional experience and honors are set forth in my *curriculum vitae*, a copy of which is attached hereto as **Tab A (Ex. 2055)**.

6. In 1980, I was Professor of Molecular Biology and the Director of the Institute of Molecular Biology at Zürich University.

7. In January 1980, Dr. Tadatsugu Taniguchi telephoned me and told me that he had cloned and was in the process of determining the complete nucleotide sequence of the human fibroblast interferon gene, and that Dr. Mark Ptashne of Harvard University had invited him to his Laboratory at Harvard University to express the cloned full length gene in *E coli*.

8. At a meeting I had with Dr. Taniguchi in Zurich, Switzerland at the end of February 1980, we decided to prepare a manuscript for publication in which the sequence similarities between leukocyte – which I and co-workers had cloned and sequenced - and fibroblast interferons was discussed. We, along with co-authors, prepared a manuscript entitled "Human Leukocyte and Fibroblast Interferons are Structurally Related" for submission for publication. Both interferon sequences appear in Figure 1. A copy of the manuscript, which I will refer to as "Taniguchi preprint" is attached hereto as **Tab B (Ex. 2023)**.

9. In late March 1980, I caused the Taniguchi preprint to be distributed to numerous researchers in the field, along with a preprint entitled "The nucleotide sequence of a cloned human leukocyte interferon cDNA," authored by Ned Mantei, myself and others, ("Mantei preprint") attached hereto as **Tab C (Ex. 2047)**. I distributed these preprint manuscripts without restrictions, and under no confidentiality obligations.

10. I have reviewed the attached copy of the "Mailing List for 'Interferon' Manuscripts" attached hereto as **Tab D (Ex. 2024)**. I created this list. The choice of persons on the list was based on my knowledge of scientists and others who were working on projects directly or indirectly related to interferons, including the cloning, structure, function, and clinical applications of interferons. I believe this copy to be a true and accurate copy of the list of scientists and others to whom the Mantei preprint and Taniguchi preprint were distributed in late March 1980.

11. The distribution list includes the name "Vilcek" of "New York." I know that this referred to Dr. Jan Vilcek, who was at New York University in late March 1980.

12. The distribution list includes the name, "Petska" of "Nutley." I know that this referred to Dr. Sidney Petska, who was with the Roche Institute of Molecular Biology in late March 1980. Dr. Petska worked at the Nutley, New Jersey location of the Roche Institute.

13. The distribution list also included individuals working for Roche at the Nutley, NJ location in March 1980 in addition to Dr. Petska, including "Ochoa", "Skalka", and "Horeckar."

14. The distribution list includes the name "Harris" of "Adelaide, S. Austr." I know that this referred to Dr. R.J. Harris, who was a Lecturer in Biochemistry at the South Australian Institute of Technology in late March 1980.

15. I have reviewed the attached copy of the April 10th, 1980 letter and attachment from R.J. Harris to myself (**Tab E; 2025**). The attachment to the letter is a copy of part of Fig. 1 from the Taniguchi preprint with markings indicating some sequence comparison analysis. The letter also has a stamp of "18. April 1980," which indicates that I most likely received the letter on April 18, 1980.

16. The Mantei preprint was published in the journal Gene: Mantei et al., Gene, 10, pp. 1-10 (1980), (**Tab F; Ex. 2049**).

17. The Taniguchi preprint was published in the journal Nature: Taniguchi et al., Nature, Vol. 285, pp. 547-549 (June 1980), (**Tab G; Ex. 2044**).

18. I have reviewed the attached copy of the Taniguchi preprint (**Tab B; Ex. 2023**) and I believe this copy to be a true and accurate copy of the Taniguchi preprint I distributed in late March 1980.

19. I have reviewed the attached copy of the Mantei preprint (**Tab C; Ex. 2047**) and I believe this copy to be a true and accurate copy of the Mantei preprint I distributed in late March 1980.

20. I reserve the right to revise, supplement, and amend this declaration.

21. I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements are made with the knowledge that willful false statements so made are punishable by fine or imprisonment or both under § 1001 of Title 18 of the United States Code and that such willful statements may jeopardize the validity or enforceability of any patent or patent application surviving this interference.

DATED: February 13, 2007 BY: _____
CHARLES WEISSMANN, M.D., PH.D.

Tab A

CURRICULUM VITAE

Professor Charles Weissmann, M.D., Ph.D., For.Mem.R.S.

Personal Information

Born October 14, 1931
Lived in Zürich from 1931-1941, 1946-1961 and 1967 to 1999; in Rio de Janeiro from 1941-1946; in New York from 1961-1967; in London from 1999-2004
Moved to Palm Beach, Florida, March 1, 2005
Swiss Citizen - Married, 4 children

Education

Kantonales Gymnasium, Zürich 1946-1950
Zürich University, 1950-1961
Degrees of M.D. (1956) and Ph.D. in Organic Chemistry (1961)

Positions

Assistant to Professor P. Karrer, Zürich University, 1960-1961
Postdoctoral Fellow, New York University School of Medicine, Department of Biochemistry, 1961-1963,
Instructor in Biochemistry, New York University School of Medicine, Department of Biochemistry, 1963-1964
Assistant Professor in Biochemistry, New York University School of Medicine, Department of Biochemistry, 1964-1965
Associate Professor in Biochemistry, New York University School of Medicine, Department of Biochemistry, 1965-1967
Professor extraordinarius in Molecular Biology, 1967-1970
Professor ordinarius in Molecular Biology, 1970-1999
Director of the Institute of Molecular Biology, University of Zürich, 1967-1999
Professor emeritus, University of Zürich, since 1999.
Senior Research Scientist and Visiting Professor, MRC Prion Unit, St.Mary's Hospital (1999-2001) and University College, London since 2001
Chairman, Department of Infectology, Scripps Florida Research Institute and Member of the Steering Committee (2004)

Awards and Honours

Ruzicka Prize in Chemistry (Switzerland, 1966)
Marcel Benoist Prize, Bern (1970)
Sir Hans Krebs Medaille, Budapest (1974)
Honorary Member of the American Society of Biological Chemistry (since 1979)
Otto Warburg Prize, Innsbruck (1980)
Member of the Deutsche Akademie der Naturforscher Leopoldina (since 1980)
Dr. H.P. Heineken Prize, Amsterdam (1982)

Sugano EXHIBIT 2055
Sugano v. Goeddel
Interference No. 105,334 and 105,337

Scheele Medal, Uppsala (1982)
Foreign Member of the Royal Society (since 1983)
Honorary Member of the American Academy of Arts and Sciences (1985)
Cancer Prize (Krebspreis) of the Schweizerische Krebsliga (1987)
Jung-Preis für Medizin, Hamburg (1988)
Foreign Associate of the U.S. National Academy of Sciences (1989)
Gabor Medal of the Royal Society (1993)
Robert-Koch Medal (1995)
Datta Lectureship Award of the FEBS (1996)
Charles-Léopold Mayer Prize of the French Academy of Science (1996)
Royal Society Glaxo Wellcome Prize (1996)
Honorary Member, Dept. of Biochemistry, University of Oxford (1997)
Member of the Schweizerische Akademie der Medizinischen Wlssenschaften (1997)
August-Wilhelm-von-Hofmann-Denkmünze (Gesellschaft Deutscher Chemiker, Wien 1997)
Klaus-Joachim-Zülch-Preis (Max-Planck-Gesellschaft, 1997)
Max Delbrück Medal (Berlin, 1997)
Wilhelm-Exner-Medaille (Wien, 1997)
Distinguished Service Award (Miami, 1998)
Corresponding Member of the Nordrhein-Westfälischen Akademie der Wissenschaften (1998)
Ausländisches Mitglied des Orden pour le mérite für Wissenschaften und Künste (Bonn, 1998)
Mendel Medal (Genetical Society, London, 1998)
Extraordinary Member of The Berlin-Brandenburgischen Akademie der Wissenschaften (Berlin, 1999)
Samuel Rudin Distinguished Visiting Professor (1999, Columbia University, N.Y.)
Fellow of the American Academy of Microbiology (Washington, 1999)
Visiting Professor, Rochester University (2001)
Visiting Professor, Imperial College of Medicine (1999-2002)
Betty and David Koetser Award (Zürich, 2001)
Fellow of the Academy of Medical Sciences (London, 2001)
Friedrich-Bauer-Prize for Medical Research (University of Münich, 2001)
Honorary Senior Fellow, Institute of Neurology, University College London (2004)
Warren Alpert Foundation Prize (Harvard Medical School, Boston, 2004)
Distinguished Research Professor, Dept. Biological Sciences, Florida Atlantic University (2004)
Fifth Annual Dart/NYU Biotechnology award (New York University, April 2006)

Honorary Degrees

Doctor *honoris causa*, University of Verona (1992)
Doctor *honoris causa*, University of Gent (1994)
Doctor *honoris causa*, ETH Zürich (1998)
Doctor *honoris causa*, University of Zürich (2000)
Doctor *honoris causa*, University of St.Andrews (St.Andrews, 2000)
Doctor *honoris causa*, Ecole Federal Polytechnique (Lausanne, 2001)

Other activities

Member of the Editorial Board of Biochimica et Biophysica Acta (1965-1968)
Associate Managing Editor of Biochimica et Biophysica Acta (1968-1980)
Member of the Editorial Board of Gene (1980-1983)
Member of the Editorial Board of the EMBO Journal (1982-1986)
Member of the European Molecular Biology Organization (EMBO) (since 1968)
Member of the Schweizerische Kommission für Molekularbiologie (SKMB) (1968-1971)
President of the Zürcher Chemische Gesellschaft (1969-1970)
President of the Schweizerische Gesellschaft für Zell- und Molekularbiologie (1970-1972)
President of the Roche Research Foundation (1971-1977)
Member of the Scientific Board of Biogen (1978-1988)
Chairman of the Scientific Board of Biogen (1984-1986)
Associate Editor of Cell (1983-1988)
Member of the Board of Governors of the Weizmann Institute of Science (since 1985)
President of the Ernst Hadorn Stiftung (since 1986)
Member of Scientific Advisory Board ZMB, Heidelberg (1988- 1990)
Member of the Scientific Council of the Swiss National Fund (1989-1994) and President of the Section IIIA (1992-1994)
Member of the Board of Directors of F. Hoffmann-La Roche Ltd., Basel (1989-2001)
Member of the Human Genome Organisation (HUGO) (since 1989)
Member of the Academia Europaea (since 1989)
Member of the International Scientific Advisory Board of the Netherlands Cancer Institute (Amsterdam)
Member of the Scientific Advisory Board of the Roche Institute of Molecular Biology, Nutley (1993-1995)
Member of the Scientific Advisory Board of the Osaka Bioscience Institute, Osaka (1993-1998)
Member of the Scientific Advisory Board of the Institut Suisse de Recherche sur le Cancer (ISREC), Lausanne (1994-1999)
Member of the Scientific Advisory Board of Roche Molecular Systems, Alameda Ca. (1994-98)
Member of the Scientific Council of the International Human Frontiers Research Program (1994-1998)
Associate Editor of Molecular Medicine (1994-2000)
Chairman of the European Commission Group on Bovine Spongiform Encephalopathy (1996)
Member of the Board of Governors of Tel Aviv University (since 1997)
Member of the Editorial Board of the Proceedings of the Royal Society (since 1999)
Member of Board of Directors of Speedel (2003-2004)
Member of the Editorial Board of the Journal of NeuroVirology (since 2006)
Member of the Editorial Board of Prion (since 2006)

Special Lectures

Sir Hans Krebs Lecture (Budapest, 1975)
Scheele Lecture (Uppsala, 1982)
Harvey Lecture (New York1982)
Severo Ochoa Lecture (Perth, 1983)
von Siemens Lecture (München, 1993)
Erna Struckmann Lecture (Heidelberg 1995)
Hertz Lectures (Tel Aviv University, 1996)
Third Carmen and Severo Ochoa Memorial Lecture (Oviedo, 1996)
Wright Lecture (Geneva, 1996)
Irving Segal Lecture (Merck, 1996)
Erste Charles-Rudolph-Brupbacher Vorlesung (Zürich, 1997)
Yehouda Levi Memorial Lecture (Tel Aviv, 1997)
Aharon Katzir Lecture (Rehovoth, 1997)
Sackler Lecture (Cambridge, 1997)
Ho-Wang Lee Lecture (Am.Soc.Virol., 1997)
Antonini Lecture (Italian Soc.Biochem., Ancona 1997)
Berlin Lecture on Molecular Medicine (Berlin1997)
Pauli Vorlesung (ETH, Zürich 1998)
Roger Sohier Lecture (IARC, Lyon, 1988)
Neurobiology Lecture 1998 (Biocenter, Basel)
Kunio Yagi IUBMB Plenary Lecture (Jerusalem, 1998)
J.F.Heremans Lecture (Univ.catholique de Louvain, 1998)
"The Anatomy Lecture" (Amsterdam, 1998)
Mendel Lecture (Genetical Society London, 1998)
Mendel Vorlesung (Wien, 1998)
Mayr Lecture (Berlin, 2000)
EMBL Distinguished Visitor Lecture (2001)
Wright Lecture (Göttingen, 2001)
Distinguished Lecture, American College of Neuropsychopharmacology, (Waikoloa, Hawaii, 2001)
Keynote Lecture, 2nd International Congress on Immunosuppression (San Diego, 2001)
Honors Lecture, New York University Medical School (New York, 2002)
Keynote Lecture, Sackler Colloquium, National Acad.Sci (Washington, March 2002)
Keynote Lecture, 3rd Pfizer Symposium (Charleston, 2002)
Israel Pollak Distinguished Lecturer (Haifa, 2003)
Keynote EMBO Lecture (St.Moritz, 2004)
Bernard Fields Lecture (TSRI, La Jolla, 2005)
Henry Kunkel Lecture (Cambridge, 2005)
Keynote Lecture, 6th International Symposium of Familial Amyloidotic Polyneuropathy (La Jolla, 2005)
Twelfth Geoffrey H. Bourne Memorial Lecture (St.George's University, Grenada, February 2006)
Keynote Lecture, International Society of NeuroVirology (Philadelphia, 2006)
Sackler Lecture (MIT, Boston, 2006)
[Inaugural Speaker Professor Severo Ochoa Lectures (NYU, New York 2007)]

Current address:

Lab:
Dept. of Infectology
Scripps Florida
5353 Parkside Drive RF2
Jupiter, Florida 33480
Direct phone: 001 561 799 8910
Office phone: 001 561 799 8895
Office fax: 001 561 799 8960
e-mail: charlesw@scripps.edu

Home:
100, Sunrise Ave., apt. 603
Florida 33480
Tel: 001 561 543 2500

TAB B

mar - apr. 1980

Human leukocyte and fibroblast interferons are structurally related.

Tadatsugu Taniguchi[+], Ned Mantei[o], Marco Schwarzstein[o], Shigekazu Nagata[o], Masami Muramatsu[+] and Charles Weissmann[o]

[o]Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

[+]Dept. of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 170, Japan

SUMMARY

The coding sequences of the cDNAs of cloned human leukocyte interferon I and human fibroblast interferon show homologies of 45% at the nucleotide and 29% at the amino acid level. We estimate that the two genes were derived from a common ancestor about 500 to 1'000 million years ago.

The acid-stable human interferons are subdivided into two major groups, namely fibroblast interferons (F-IF) and leukocyte interferons (Le-IF); these are the major components of the interferons produced by induced fibroblasts and leukocytes, respectively. Some cells, such as the lymphoblastoid Namalva cell line, produce a mixture of 90% Le-IF and 10% F-IF (1, 2). The two interferon types have several features in common: both are glycoproteins with molecular weights ranging from 16'000 to 26'000 (3-9), the induction and shut-off of their synthesis appears to be under similar control (6), and at least some of the responses elicited in target cells are similar, such as induction of an antiviral state, which is accompanied by increased synthesis of several proteins (10-13). Nonetheless, the two kinds of interferons differ in many respects. Antibodies directed against Le-IF do not neutralize F-IF and vice-versa (14), the target cell specificities of the two IFs differ (15), and the sequences of the 13 amino terminal amino acids of F-IF and of Le-IF (from lymphoblastoid cells) show no homology (16, 22). Although Le-IF and F-IF are encoded by different mRNA species (17), it is not known whether these mRNAs are transcribed from distinct genes or whether they arise from the same gene via a common precursor which is processed or spliced in different modes.

We have recently cloned and sequenced one species each of Le-IF (Le-IF I) (18,19) and F-IF cDNA (20,21). A second

species of Le-IF (Le-IF II) cDNA has recently been identified (M. Streuli, S. Nagata and C. Weissmann, unpublished results).

In Fig. 1 the nucleotide sequences of Le-IF I and F-IF cDNA were aligned so that the AUGs closest to their 5' termini coincided. From the amino terminal sequence published for F-IF (16) and lymphoblastoid Le-IF (22) one can determine that in the case of F-IF the 21st codon following the initiation triplet and in the case of Le-IF the 23rd codon represents the first amino acid of the interferon polypeptide. Presumably the stretch in between encodes a signal peptide. Since the putative signal peptide of Le-IF comprises 23 and that of F-IF 21 amino acids, the IF polypeptides, as aligned in Fig. 1, are shifted by two residues relative to their termini. In this alignment, 48 of 166 positions (29%) of the interferon polypeptides have identical amino acids. To plot the degree of homology between the F-IF and Le-IF as function of the map distance, the sequence was subdivided into segments of 8 amino acids (or 24 nucleotides), each overlapping by 4 amino acids (or 12 nucleotides) with the neighboring segments, and the percent coincidence of amino acids (and nucleotides) for each segment was determined (cf. van Ooyen et al., ref. 23). As seen in Fig. 2, amino acid sequences show three domains of homology. The first one, with the least degree of homology, corresponds to the putative signal sequence, which is rich in hydrophobic residues and has 4 identical amino acid positions out of 21; the second domain, between the 28th and 80th amino acid (counted on the Le-IF sequence), has 21 identical residues out of 51 (41% homology) and the third, between

positions 115 and 151 (Le-IF sequence), has 19 out of 35 identical residues (54%). The longest stretches of contiguous conserved amino acids are Gln-Phe-Gln-Lys (pos. 47-50 of Le-IF and 49-52 of F-IF) and Cys-Ala-Trp (pos. 139-141 and pos. 141-143, respective the latter sequence is notable because it comprises Cys and Trp, which are preferentially conserved in related proteins (24). Table 1 shows that conservation was highest between the interferon polypeptides (not considering the signal sequences) for Trp, Phe, Arg, Cys and Tyr residues, in agreement with the general experience that the amino acids most likely to be conserved between related proteins are Trp > Cys > Tyr > Arg > Phe, His (24). Even where amino acids are conserved, the codons show one or more nucleotide changes in half the instances. The codons of 3 out of 7 conserved Leu residues are non-related, as are 2 of 4 codons pertaining to conserved Ser residues. This suggests that there is a strong selective pressure favoring the conservation of several amino acids. It is quite likely that at least some of the conserved amino acids are essential for a function common to Le-IF and F-IF, perhaps the induction of the virus-resistant state in the target cell. These findings may provide guidelines for the tailoring of modified (25), possibly shorter polypeptides possessing certain activities of interferon.

The nucleic acid sequences show an average homology of 43% in the domain of the signal sequence and of 45% in the interferon polypeptide sequence. On a random basis, about 25% of the nucleotide positions should

coincide. Within the interferon coding sequence, the nucleotide homologies are more evenly distributed than the amino acid homologies. However, one may distinguish, albeit to a less pronounced degree, the same two blocks of similarity noted for the amino acids. The longest region without mismatches extends for 13 nucleotides (cf. 47th to 51st codon of Le-IF vs. 49th to 53rd codon of F-IF). There are, in addition, sequences of 17, 18 and 20 nucleotides with 3, 3 and 4 mismatches, respectively. The heteropolymeric 3'-terminal non-coding region of Le-IF cDNA has 242 nucleotides, and is longer by 39 residues than its counterpart in F-IF cDNA. In aligning the two sequences four gaps were introduced to maximize homology, as described by van Ooyen et al.[23] Thereby, several segments were matched with 29 to 41% homology. The introduction of gaps in the alignment may be justified in view of the arguments presented previously, that introns and non-coding regions of reduplicated genes diverge as a consequence of block insertions and/or deletions in the course of evolution (23,26).

It is unlikely that the extent of homology between Le and F-IF cDNA would allow meaningful crosshybridization between the two species.

On the basis of our findings there is no doubt that Le-IF and F-IF genes are derived from a common ancestral sequence. When did the separation of these genes occur? Human α and β globin show 57% amino acid mismatches, and human β-globin and myoglobin, as well as α-globin and myoglobin, 76% mismatches.

If the rate of divergence of interferons and globins is comparable (however, cf. p. 50, ref. 24, for proteins showing both higher and lower rates) then the separation of interferon genes occurred after that of myoglobin and hemoglobins and before that of α- and β-globins, i.e. between 500 and 1000 million years ago (24). The interferon genes may thus be about as old as the vertebrates (27).

ACKNOWLEDGEMENTS

T.T. and M.M. are indebted to Dr. H. Sugano for his continuous support and interest. The work at the University of Zürich was supported by Biogen S.A. and the Schweizerische Nationalfonds.

REFERENCES

1. Havell, E.A., Yip, Y.K. & Vilček, J. J. gen. Virol. 38, 51-59 (1977).
2. Paucker, K. The Interferon System, Texas Reports on Biology and Medicine (Baron, S. & Dianzani, F., eds.) 35, 23-28 (1977).
3. Bridgen, P.J., Anfinsen, C.B., Corley, L., Bose, S., Zoon, K.C., Rüegg, U.T. & Buckler, C.E. J. Biol. Chem. 252, 6585-6587 (1977).
4. Knight, E., jr. Proc. Natl. Acad. Sci. USA 73, 520-523 (1976).
5. Tan et al., J. Biol. Chem. 254, 8067-8073 (1979).
6. Hayes, T.G., Yip, Y.K. & Vilček, J. Virology 98, 351-363 (1979).
7. Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. & Pestka, S. Proc. Natl. Acad. Sci. USA 76, 640-644 (1979).
8. Stewart, W.E., II, & Wiranowska-Stewart, M. Virology 97, 473-476 (1979).
9. Havell, E.A., Yamazaki, S. & Vilček, J. J. Biol. Chem. 252, 4425-4427 (1977).
10. Lebleu, B., Sen, G.C., Shaila, S., Cabrer, B. & Lengyel, P. Proc. Natl. Acad. Sci. USA 73, 3107-3111 (1976).
11. Hovanessian, A.G., Brown, R.E. & Kerr, I.M. Nature 268, 537-540 (1977).
12. Hovanessian, A.G. & Kerr, I.M. Eur. J. Biochem. 93, 515-526 (1979).
13. Schmidt, A., Chernajovsky, Y., Shulman, L., Federman, P., Berissi, H. & Revel, M. Proc. Natl. Acad. Sci. USA 76, 4788-4792 (1979).

14. Havell, E.A., Berman, B., Ogburn, C.A., Berg, K., Paucker, K. & Vilček, J. Proc. Natl. Acad. Sci. USA 72, 2185-2187 (1975).

15. Stewart, W.E., II, In "The Interferon System" (Springer-Verlag, Wien & New York), pp. 134-145 (1979).

16. Knight, E., jr., Hunkapiller, M.W., Korant, B.D., Hardy, R.W.F. & Hood, L.E. Science 207, 525-526 (1980).

17. Cavalieri, R.L., Havell, E.A., Vilček, J. & Pestka, S. Proc. Natl. Acad. Sci. USA 74, 3287-3291 (1977).

18. Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. & Weissmann, C. Nature, in press (1980).

19. Mantei, N., Schwarzstein, M., Streuli, M., Panem, S., Nagata, S. & Weissmann, C., Gene, in press (1980).

20. Taniguchi, T., Sakai, M., Fujii-Kuriyama,Y., Muramatsu, M., Kobayashi, S. & Sudo, T. Proc. Jap. Acad. Ser. B, 55, 464-469 (1979).

21. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. Gene, in press (1980).

22. Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapiller, M.W. & Hood, L.E. Science 207, 527-528 (1980).

23. van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. Science 206, 337-344 (1979).

24. Dayhoff, M.O., In "Atlas of Protein Sequence and Structure" (Natl. Biomedical Res. Found., Washington, D.C.) Vol. V (1972

25. Müller, W., Weber, H., Meyer, F. & Weissmann, C. J. Mol. Biol 124, 343-358 (1978).

26. Konkel, D.A., Maizel, J.V., jr. & Leder, P. Cell 18, 865-873 (1979).

27. Kurtén, B. In "Ideas in modern biology" Proceedings vol. 6, XVI International Congress of Zoology (Moore, J.A., ed.), Natural History Press, N.Y. (1965) pp. 327-354.

TABLE 1 Conservation of amino acids in leukocyte and fibroblast interferon.*

	F-IF	Le-IF	Conserved amino acids	Number of changes in codon of conserved amino acids 0	1	2	3
Leu	25	22	8	1	4	3	
Cys	3	5	2	1	1		
AsN	12	6	1	1			
Arg	11	12	5	1	3	1	
Phe	9	8	4	2	2		
Pro	1	6	1		1		
GlN	11	10	3	3			
Lys	11	8	3	2	1		
Ala	6	10	2	2			
Glu	13	15	4	4			
Ile	11	7	3	2	1		
Ser	9	13	4		2	1	1
Trp	3	2	2	2			
Tyr	10	4	4	1	3		
Val	5	6	1	1			
Asp	5	11	1	1			
Thr	6	9	0				
Gly	6	3	0				
Met	4	6	0				
His	5	3	0				
	166	166	48	24	18	5	1

*The data are from Taniguchi et al. (ref. 21) and Mantei et al. (ref. 19).

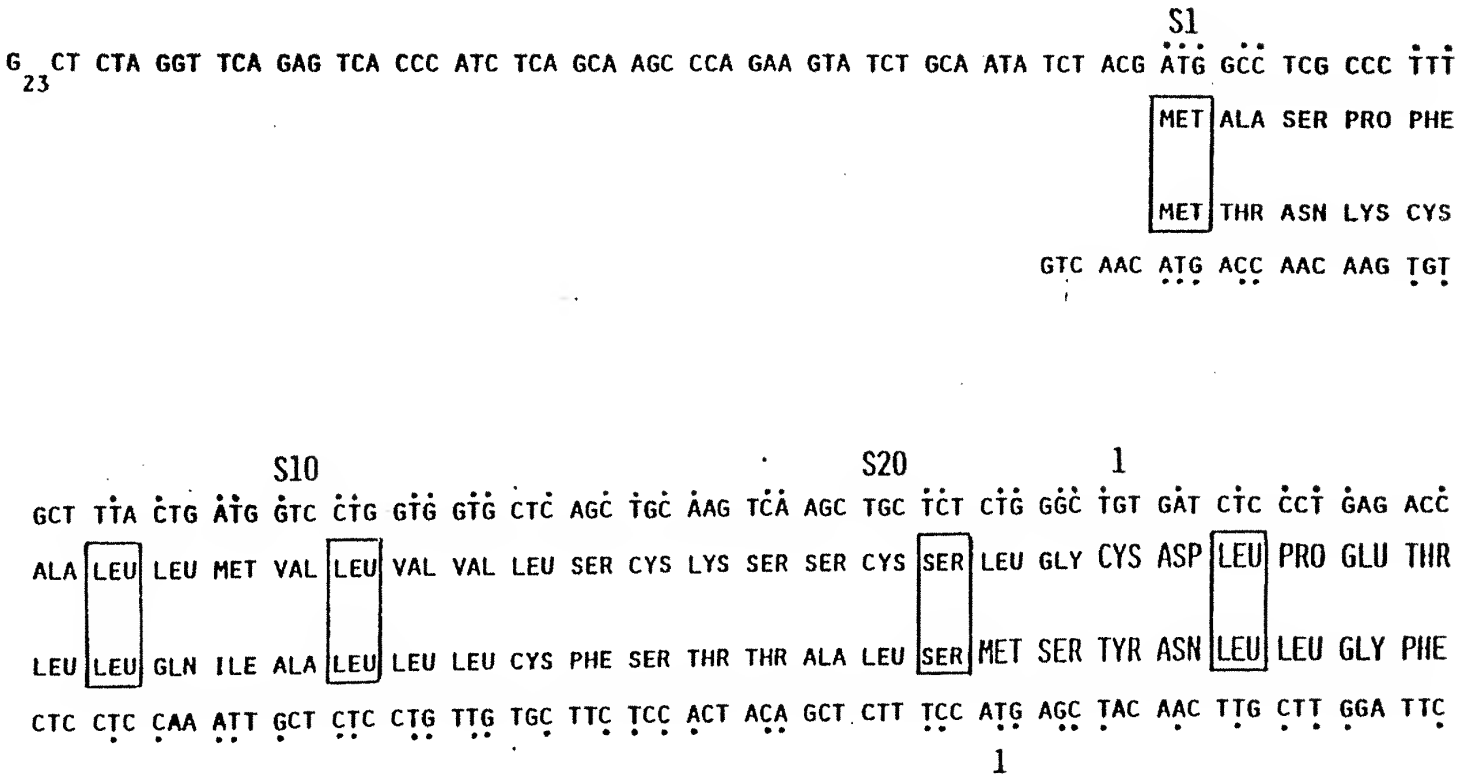
FIGURE LEGENDS

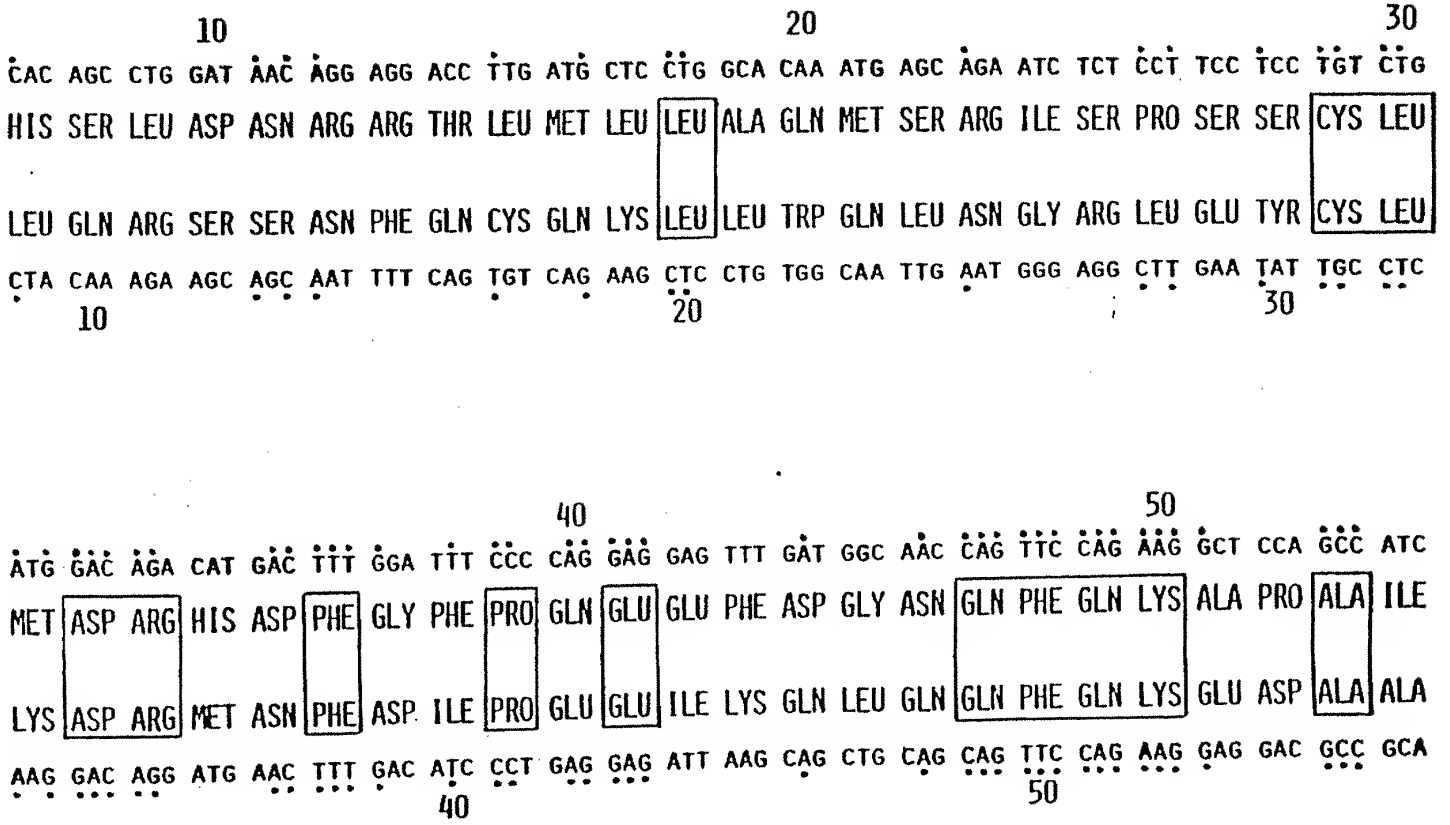
Fig. 1 Comparison of the nucleotide sequences of human leukocyte interferon I (Le-IF I) and human fibroblast interferon cDNA and of the derived amino acid sequences. The sequences are from Mantei et al.(19) and Taniguchi et al. (21). They were aligned to give maximal homology. Identical amino acids are framed, identical nucleotides are marked by a dot. S1 to S23 indicate the amino acids of the putative signal sequence; 1 to 166 the amino acids of the interferon polypeptides.

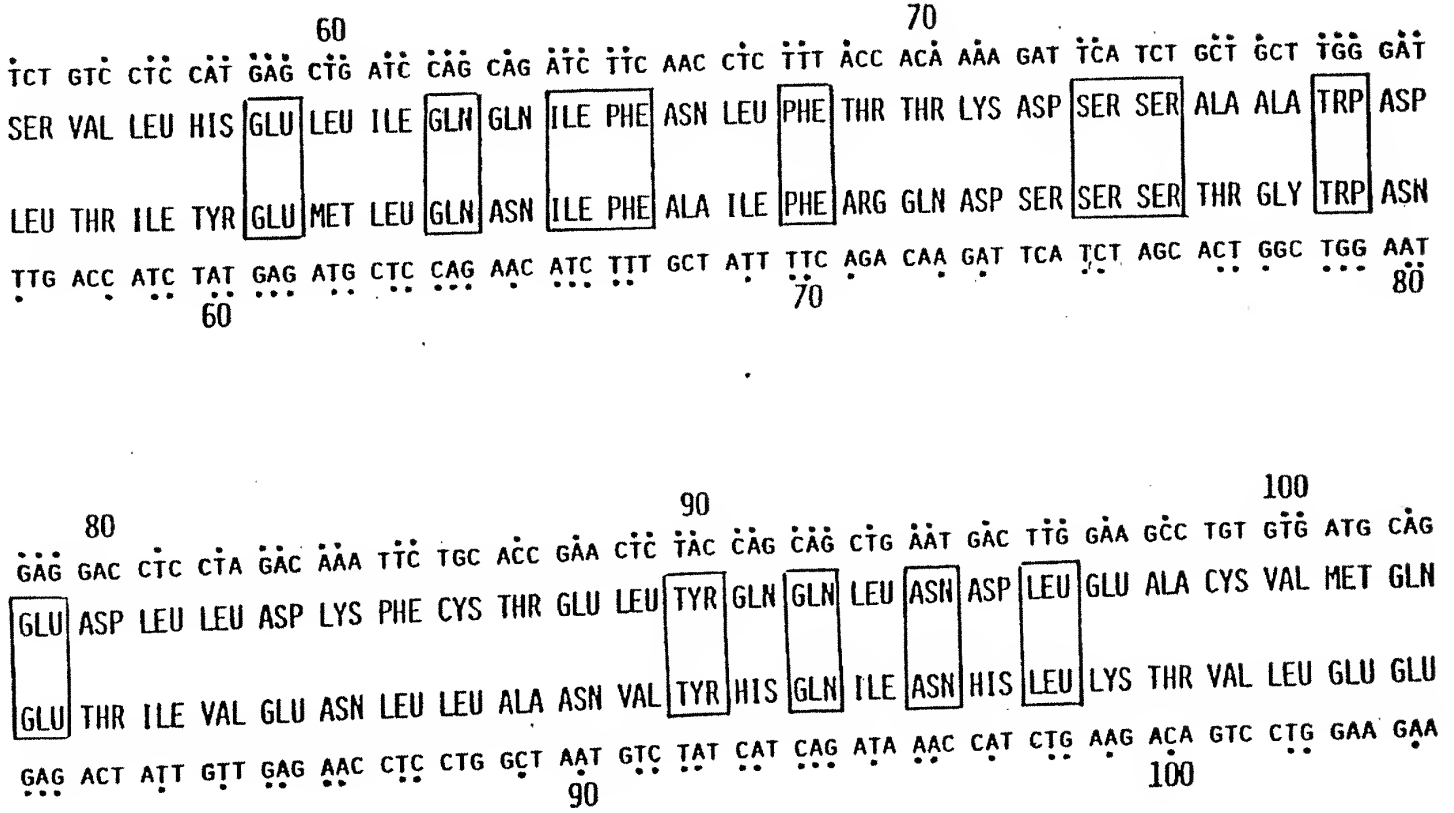
Fig. 2 Similarity of the nucleotide and amino acid sequences of human leukocyte interferon I and fibroblast interferon.

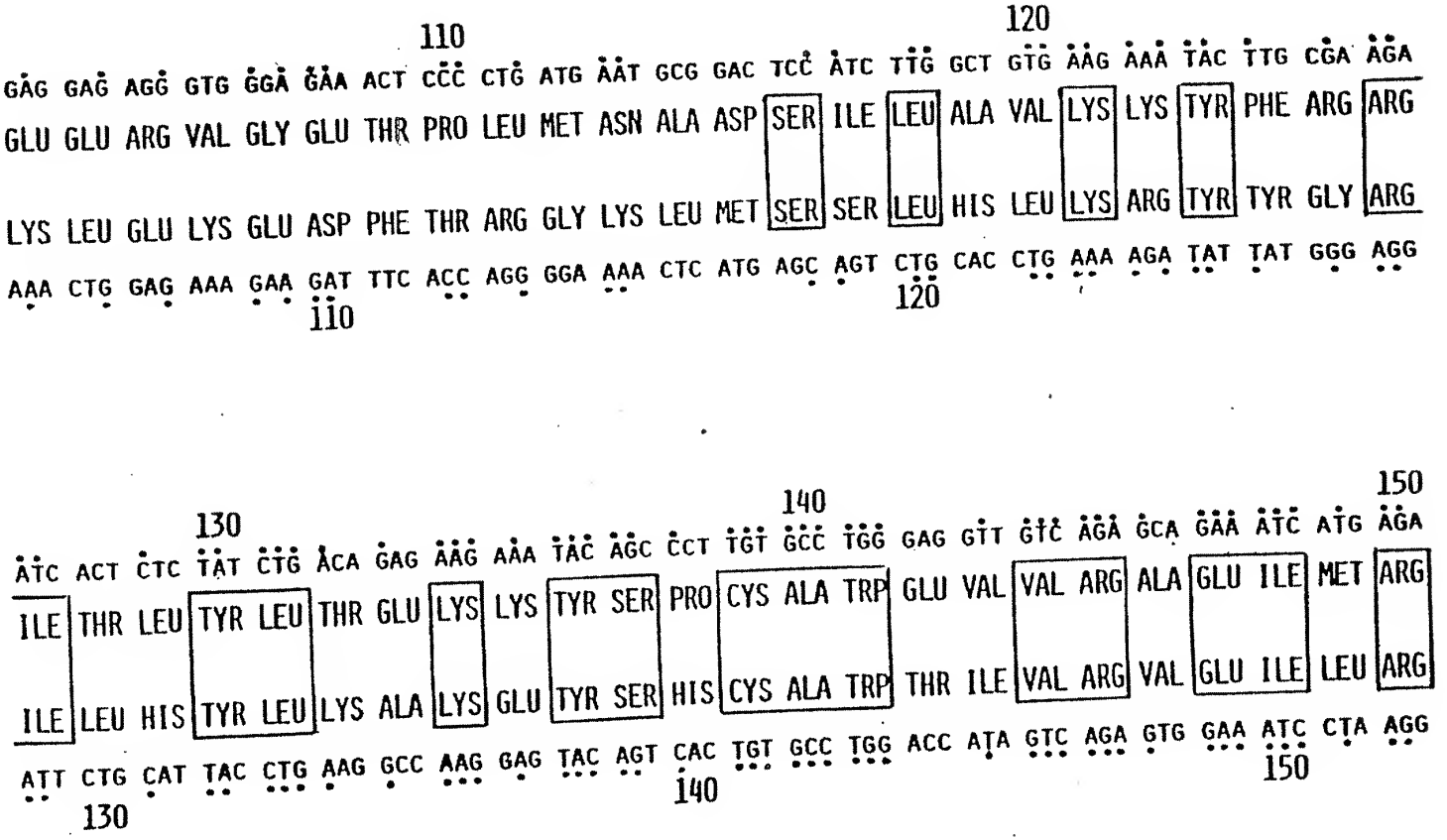
The sequences shown in Fig. 1 were subdivided in segments of 8 amino acids or 24 nucleotides, each overlapping by 4 and 12 residues, respectively, with the neighboring segments. The percentage of coincident residues was plotted as a function of map position. Open vertical blocks, nucleotides; filled vertical blocks, amino acids. L-IF, leukocyte interferon cDNA; F-IF, fibroblast interferon cDNA; lines, non-coding sequences; hatched bars, putative signal peptide; open bars, interferon polypeptide.

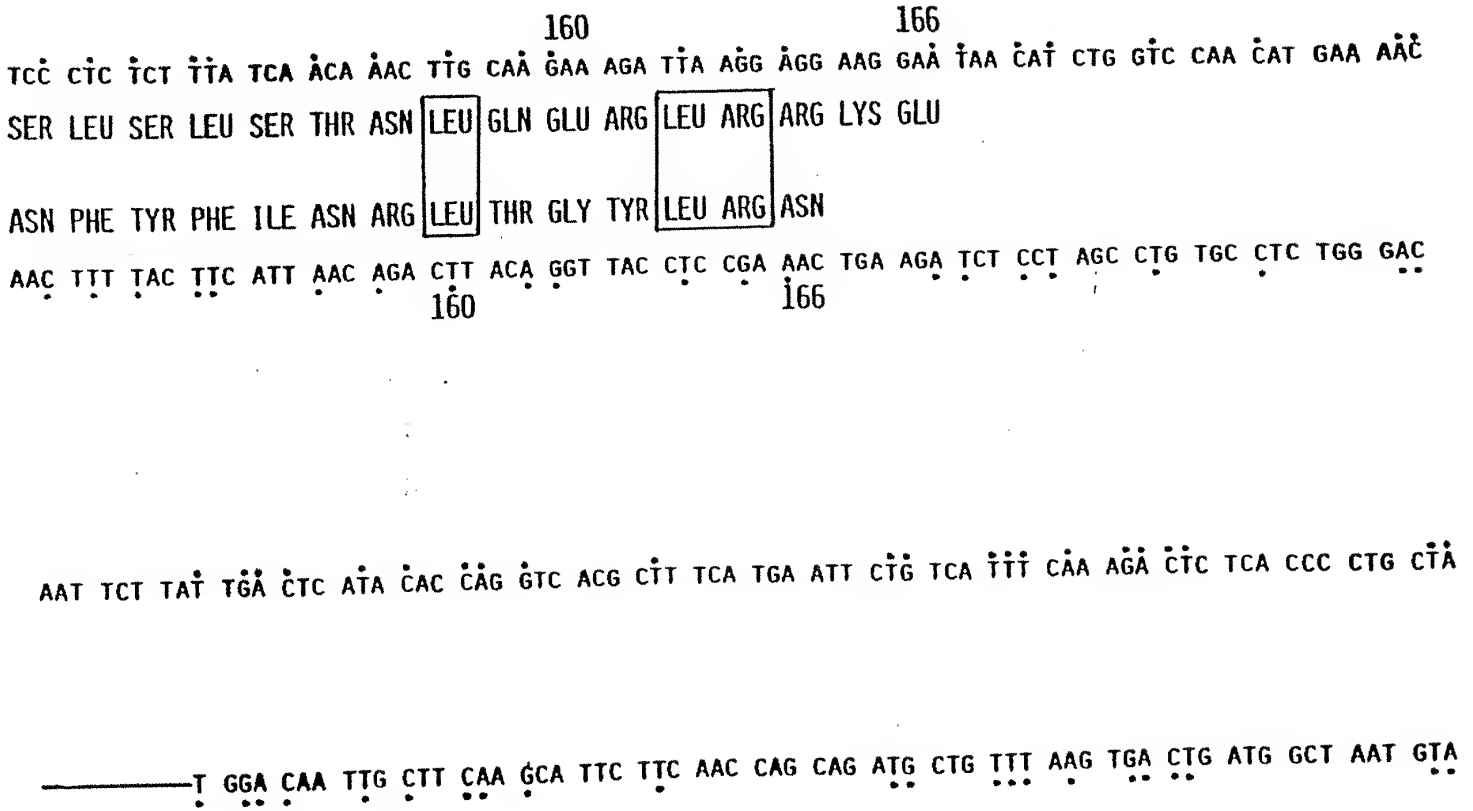
Fig. 1


S1
G CT CTA GGT TCA GAG TCA CCC ATC TCA GCA AGC CCA GAA GTA TCT GCA ATA TCT ACG ATG GCC TCG CCC TTT
23
MET ALA SER PRO PHE
MET THR ASN LYS CYS
GTC AAC ATG ACC AAC AAG TGT
S10
S20
1
GCT TTA CTG ATG GTC CTG GTG GTG CTC AGC TGC AAG TCA AGC TGC TCT CTG GGC TGT GAT CTC CCT GAG ACC
ALA LEU LEU MET VAL LEU VAL VAL LEU SER CYS LYS SER SER CYS SER LEU GLY CYS ASP LEU PRO GLU THR
LEU LEU GLN ILE ALA LEU LEU LEU CYS PHE SER THR THR ALA LEU SER MET SER TYR ASN LEU LEU GLY PHE
CTC CTC CAA ATT GCT CTC CTG TTG TGC TTC TCC ACT ACA GCT CTT TCC ATG AGC TAC AAC TTG CTT GGA TTC
1


10
20
30
CAC AGC CTG GAT AAC AGG AGG ACC TTG ATG CTC CTG GCA CAA ATG AGC AGA ATC TCT CCT TCC TCC TGT CTG
HIS SER LEU ASP ASN ARG ARG THR LEU MET LEU LEU ALA GLN MET SER ARG ILE SER PRO SER SER CYS LEU
LEU GLN ARG SER SER ASN PHE GLN CYS GLN LYS LEU LEU TRP GLN LEU ASN GLY ARG LEU GLU TYR CYS LEU
CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA TAT TGC CTC
10
20
30
40
50
ATG GAC AGA CAT GAC TTT GGA TTT CCC CAG GAG GAG TTT GAT GGC AAC CAG TTC CAG AAG GCT CCA GCC ATC
MET ASP ARG HIS ASP PHE GLY PHE PRO GLN GLU GLU PHE ASP GLY ASN GLN PHE GLN LYS ALA PRO ALA ILE
LYS ASP ARG MET ASN PHE ASP ILE PRO GLU GLU ILE LYS GLN LEU GLN GLN PHE GLN LYS GLU ASP ALA ALA
AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC GCC GCA
40
50


60
70
TCT GTC CTC CAT GAG CTG ATC CAG CAG ATC TTC AAC CTC TTT ACC ACA AAA GAT TCA TCT GCT GCT TGG GAT
SER VAL LEU HIS GLU LEU ILE GLN GLN ILE PHE ASN LEU PHE THR THR LYS ASP SER SER ALA ALA TRP ASP
LEU THR ILE TYR GLU MET LEU GLN ASN ILE PHE ALA ILE PHE ARG GLN ASP SER SER SER THR GLY TRP ASN
TTG ACC ATC TAT GAG ATG CTC CAG AAC ATC TTT GCT ATT TTC AGA CAA GAT TCA TCT AGC ACT GGC TGG AAT
60
70
80
80
90
100
GAG GAC CTC CTA GAC AAA TTC TGC ACC GAA CTC TAC CAG CAG CTG AAT GAC TTG GAA GCC TGT GTG ATG CAG
GLU ASP LEU LEU ASP LYS PHE CYS THR GLU LEU TYR GLN GLN LEU ASN ASP LEU GLU ALA CYS VAL MET GLN
GLU THR ILE VAL GLU ASN LEU LEU ALA ASN VAL TYR HIS GLN ILE ASN HIS LEU LYS THR VAL LEU GLU GLU
GAG ACT ATT GTT GAG AAC CTC CTG GCT AAT GTC TAT CAT CAG ATA AAC CAT CTG AAG ACA GTC CTG GAA GAA
90
100


110
120
GAG GAG AGG GTG GGA GAA ACT CCC CTG ATG AAT GCG GAC TCC ATC TTG GCT GTG AAG AAA TAC TTG CGA AGA
GLU GLU ARG VAL GLY GLU THR PRO LEU MET ASN ALA ASP SER ILE LEU ALA VAL LYS LYS TYR PHE ARG ARG
LYS LEU GLU LYS GLU ASP PHE THR ARG GLY LYS LEU MET SER SER LEU HIS LEU LYS ARG TYR TYR GLY ARG
AAA CTG GAG AAA GAA GAT TTC ACC AGG GGA AAA CTC ATG AGC AGT CTG CAC CTG AAA AGA TAT TAT GGG AGG
110
120
130
140
150
ATC ACT CTC TAT CTG ACA GAG AAG AAA TAC AGC CCT TGT GCC TGG GAG GTT GTC AGA GCA GAA ATC ATG AGA
ILE THR LEU TYR LEU THR GLU LYS LYS TYR SER PRO CYS ALA TRP GLU VAL VAL ARG ALA GLU ILE MET ARG
ILE LEU HIS TYR LEU LYS ALA LYS GLU TYR SER HIS CYS ALA TRP THR ILE VAL ARG VAL GLU ILE LEU ARG
ATT CTG CAT TAC CTG AAG GCC AAG GAG TAC AGT CAC TGT GCC TGG ACC ATA GTC AGA GTG GAA ATC CTA AGG
130
140
150

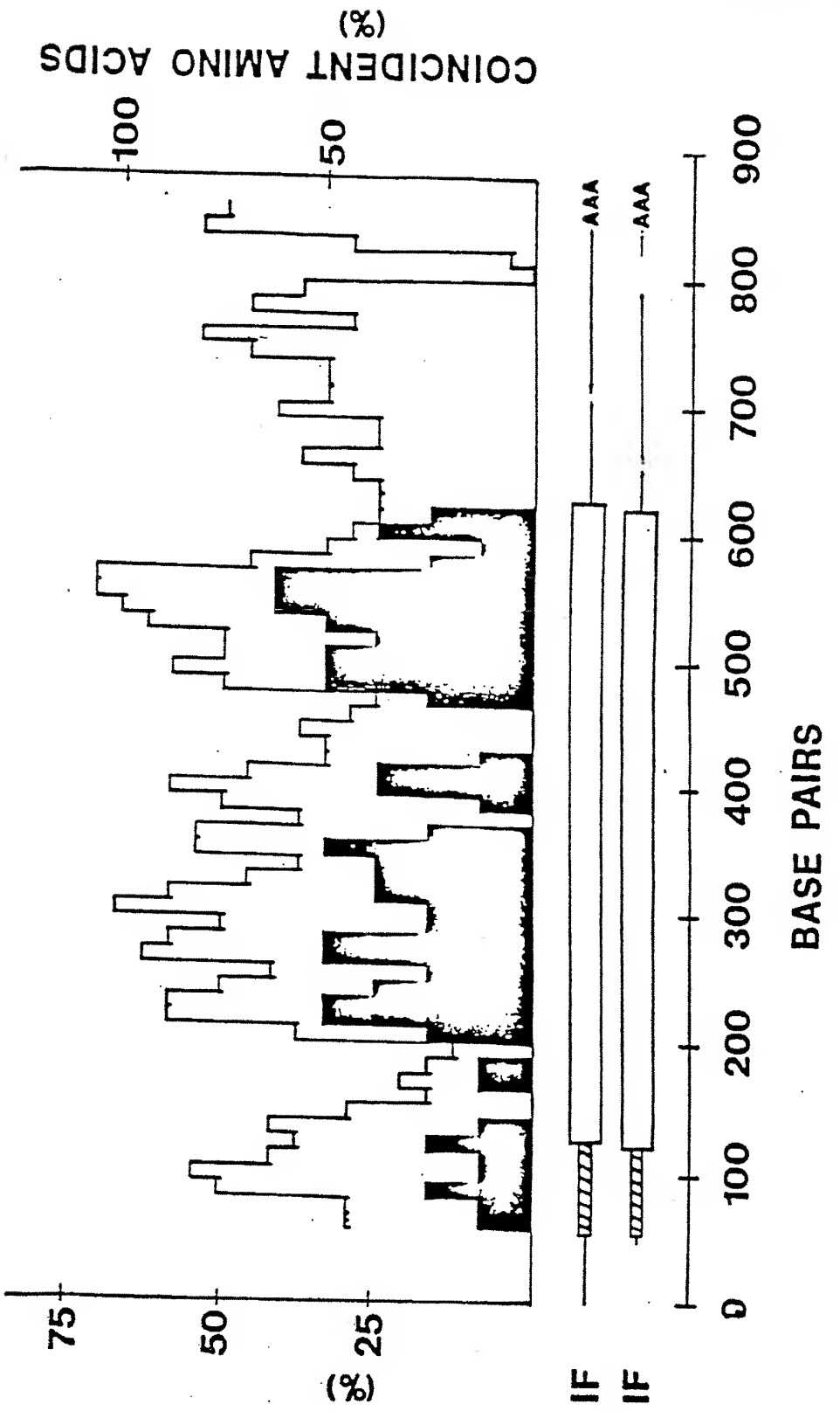

160
166
TCC CTC TCT TTA TCA ACA AAC TTG CAA GAA AGA TTA AGG AGG AAG GAA TAA CAT CTG GTC CAA CAT GAA AAC
SER LEU SER LEU SER THR ASN LEU GLN GLU ARG LEU ARG ARG LYS GLU
ASN PHE TYR PHE ILE ASN ARG LEU THR GLY TYR LEU ARG ASN
AAC TTT TAC TTC ATT AAC AGA CTT ACA GGT TAC CTC CGA AAC TGA AGA TCT CCT AGC CTG TGC CTC TGG GAC
160
166
AAT TCT TAT TGA CTC ATA CAC CAG GTC ACG CTT TCA TGA ATT CTG TCA TTT CAA AGA CTC TCA CCC CTG CTA
T GGA CAA TTG CTT CAA GCA TTC TTC AAC CAG CAG ATG CTG TTT AAG TGA CTG ATG GCT AAT GTA

-TA ACT ATG ACC ATG CTG ATA AAC TGA TTT ATC TAT TTA AAT ATT TAT TTA ACT ATT CAT AAG ATT TAA ATT

CTG CAT ATG AAA GGA CAC TAG AAG ATT TTG AAA TTT TTA TTA AAT TAT GAG TTA TTT TTA TTT ATT TAA ATT

ATT TTT GTT CAT ATA ACG TCA TGT GCA CCT TTA CAC TGT GGT TAG TGT AAT AAA ACA TGT TCC TTA TAT TTA

TT ATT TTG GAA AAT AAA TTA TTT TTG GTG CAA A-A


COINCIDENT AMINO ACIDS
(%)
100
50
75
50
25
(%)
AAA
AAA
IF
IF
0
100
200
300
400
500
600
700
800
900
BASE PAIRS

TAB C

The nucleotide sequence of a cloned human leukocyte interferon cDNA.

(Amino acid sequence; restriction map; signal sequence; interferon synthesis)

Ned Mantei, Marco Schwarzstein, Michel Streuli, Sandra Panem*, Shigekazu Nagata and Charles Weissmann

Institut für Molekularbiologie I, Universität Zürich, Hönggerberg, 8093 Zürich, Switzerland

* Present address: Dept. of Pathology, University of Chicago, Chicago, Ill. 60637, U.S.A.

This is EXHIBIT HALEY- 2
to
the Affidavit of James F. Haley, Jr.
sworn before me
this 21st day of November, 2001

Commissioner for Oath or Notary Public

MARIA E. MARLEY
Notary Public, State Of New York
No. 01MA4890484
Qualified In New York County
Commission Expires April 27, 2003

SUMMARY

We have determined the nucleotide sequence of the human leukocyte interferon cDNA-containing hybrid plasmid Z-pBR322(Pst)/HcIF-2h, which has been shown to direct the formation of a polypeptide with human leukocyte interferon activity (Nagata et al., Nature, 1980, in press). The 910 base pair insert contains a 567 (or 543) base pair coding sequence, which determines a putative preinterferon polypeptide consisting of a signal peptide of 23 (or less likely 15) amino acids, followed by an interferon polypeptide of 166 amino acids (calculated molecular weight, 19'390). The coding sequence is preceded by a (most likely incomplete) 56 bp leader and followed by a 242 bp trailer and 7 A residues from the poly(A) tail. A comparison of the sequence of 35 amino terminal amino acids of lymphoblastoid interferon (Zoon et al., Science 207, 527-528, 1980; M. Hunkapiller and L. Hood, personal communication) and the corresponding sequence deduced for leukocyte interferon revealed 9 differences. This suggests that these two interferons are encoded by two non-allelic genes.

INTRODUCTION

We have recently described the isolation of a hybrid plasmid Z-pBR322(Pst)/HcIF-2h, or Hif-2h for short, which contains a cDNA sequence coding for human leukocyte interferon. The hybrid DNA was identified by its capacity to (a) hybridize with human leukocyte interferon mRNA and (b) to direct the synthesis in E.coli of a protein with properties of human leukocyte interferon (Nagata et al., 1980).

In this paper we report the nucleotide sequence of the 910 bp insert of Hif-2h. Two AUG triplets and a UAA termination codon, all in the same reading frame, define a stretch of 507 or 543 nucleotides which encodes a polypeptide of 166 amino acids corresponding to the interferon polypeptide proper, preceded by 23 or 15 amino acids, which may constitute a signal sequence. The coding region is flanked at the 5' end by 79 nucleotides, 23 of which are terminal G residues, and at the 3' end by 264 nucleotides, 15 of which are terminal C residues.

MATERIALS AND METHODS

Plasmid DNA was prepared by method B described in Wilkie et al. (1979). EcoRI was a gift from W. Boll and BspI from A. Kiss. All other restriction enzymes were purchased from New England Biolabs and used in essence as recommended by the supplier (except that 200 µg/ml gelatin replaced bovine serum albumen in the enzyme buffers). Carrier-free [γ-^{32}P] ATP was prepared by an unpublished procedure of B. Seed.

5'-terminal labeling of DNA.

Restricted DNA (20 µg) was extracted with phenol, precipitated with ethanol, dissolved in 0.05 M Tris-HCl (pH 8), and passed over a small column of Chelex-100. Fragments with flush or 5'-overhanging ends were dephosphorylated by treatment with 0.2 units calf intestinal alkaline phosphatase (Boehringer) per pmol DNA 5' ends in 200 µl 0.05 M Tris-HCl (pH 8) for 60 min at 37°C. The enzyme was inactivated by heating 60 min at 65°C. For DNA fragments with 3' overhanging ends, bacterial alkaline phosphatase (Worthington) was used as described (Maxam and Gilbert, 1977), except that incubation was at 65°C for 30 min. The dephosphorylated DNA was purified by adsorption to and elution from DEAE-cellulose as described (Müller et al., 1978) or subjected to polyacrylamide gel electrophoresis where required (see below). Fragments recovered from a polyacrylamide (or agarose) gel in 0.15 M NaCl, 0.05 M Tris-HCl (pH 8) were adsorbed to a 0.1-ml hydroxyapatite (Biorad HTP) column, washed with 4 times 1 ml of 0.1 M potassium phosphate buffer (pH 7) and eluted with 0.3 ml 1 M potassium phosphate buffer (pH 7). The solution was diluted tenfold and the DNA adsorbed to DEAE cellulose and recovered as described (Müller et al., 1978).

After ethanol precipitation, the DNA was 5'-terminally labeled with [γ-^{32}P] ATP (12-34 µCi per pmol DNA end) and polynucleotide kinase (New England Biolabs or P-L Biochemicals Inc.) essentially as described (Maxam and Gilbert, 1977), except that the DNA

was not denatured before the kinase reaction. Specific activities of 1-1.5 μCi [^{32}P] phosphate per pmol DNA 5'-ends were obtained.

Nucleotide sequence determination.

For sequencing, labeled fragments were cleaved with a second restriction enzyme and the products separated by electrophoresis through a 5% polyacrylamide gel in Tris-borate-EDTA buffer. The desired fragments were extracted from the gel and purified as described (Müller et al., 1978). The various fragments for sequencing were prepared as follows (the number indicates the nominal fragment chain length in base pairs, the labeled site is indicated by an asterisk, and the letters in parentheses refer to the arrows shown in Fig. 1): (a) and (b), cleavage of Hif-2h with BspI, isolation by 5% polyacrylamide gel electrophoresis in Loening's buffer (Loening, 1967) of Bsp-Bsp-232 (for (a)) and Bsp-Bsp-949 (for (b)), labeling, cleavage with PstI, isolation of (a) Bsp*-Pst-83 and (b) Bsp*-Pst-827. (c) and (d), cleavage of Hif-2h with BglII, labeling, cleavage with PstI, isolation of (c) Bgl*-Pst-336 and (d) Bgl*-Pst-570. (e) and (f), cleavage of Hif-2h with MboII, labeling, digestion with PstI and HindII (to cleave an interfering 350 bp pBR322 fragment), isolation of (e) Mbo*-Pst-519 and (f) Mbo*-Pst-351. (g) and (h), cleavage of Hif-2h with EcoRI, labeling, cleavage with PstI, isolation of (g) Eco*-Pst-708 and (h) Eco*-Pst-198. (i) and (j), cleavage of Hif-2h with PstI, labeling, cleavage with BglII, isolation of (i) Pst*-Bgl-570 and (j) Pst*-Bgl-336. (k) and (l), cleavage of Hif-2h with AvaII, labeling, cleavage with PstI and BglII, isolation of (k) Ava*-Pst-186 and (l)

Ava*-Bgl-147. (m) Cleavage of plasmid with PvuII, labeling, cleavage with PstI and BglII, isolation of Pvu*-Pst-486. The fragments were degraded according to Maxam and Gilbert (1977), with the modifications described in protocols provided by the same authors in September, 1978. The products were fractionated on 0.1 x 25 x 36 cm 12% polyacrylamide gels (acrylamide/bis-acrylamide = 18/1) in 50 mM Tris-borate, 1 mM EDTA (pH 8.3), with runs of 2, 8, 18 and 26 h at 900 V following a 6 h prerun at 700 V. Best results were obtained when the gels were kept at room temperature 2-3 days before use.

RESULTS

1) Physical map of Hif-2h DNA.

Hif-2h consists of dC-elongated human Le IF cDNA joined to pBR322 (Bolivar et al., 1977) which had been cleaved with PstI and elongated with dG residues. A physical map was prepared by measuring the lengths of the fragments generated by single cleavage with EcoRI, BspI, PstI and MboII and double cleavages with PstI on the one hand and EcoRI, BglII, BspI and MboII on the other, as well as with EcoRI and BglII, and EcoRI and MboII. In addition, DNA fragments which were ^{32}P-labeled at one 5' end, were partially digested with a variety of restriction enzymes, and the lengths of the labeled products determined (Smith and Birnstiel, 1976). The resulting preliminary map was used as a basis for the nucleotide sequence analysis; the map shown in

Fig. 1 was refined (cf. also Fig. 3) using the results of the nucleotide sequence analysis described below. No restriction targets for BglI, KpnI, HaeII, XhoI, PvuI, XbaI, PstI, BstEll, BamHI, HindII, SalI, HindIII, HpaII, TaqI, HgaI, TacI, HpaI or HhaI were found in the insert; there were single sites for BspI, BglII and EcoRI, and two sites for PvuII. One each of 4 AvaII and 4 MboII targets (marked with * in Fig. 3) was not cleavable by the cognate enzyme, perhaps because of methylated bases in adjacent EcoRII and MboI sites. The MboI sites were not cleavable.

2) The orientation of the coding sequence.

In order to determine the orientation of the coding strand relative to pBR322 the experiment outlined in Fig. 2a was carried out. The hybrid plasmid Hif-2h was cleaved at the single BglII site, 5'-terminally labeled with $[^{32}P]$phosphate and digested with PstI to yield 336/344 and 578/570 bp radioactive fragments. The fragments were denatured, annealed with poly(A) RNA from induced leukocytes, and the mixture was treated with S_1 nuclease. The resulting products were denatured and analyzed by polyacrylamide gel electrophoresis. As shown in Fig. 2b, a ^{32}P-labeled fragment of about 340 nucleotides was protected. In a second, similar experiment the labeled fragments were first separated and then annealed individually with poly(A) RNA: the shorter, but not the longer probe was protected against S_1 nuclease (data not shown). These experiments identify the 5' labeled 344 nucleotide strand as the minus strand, i.e. the strand complementary

to the mRNA. Therefore, the orientation of the insert is such that the coding strand of the IF cDNA is a continuation of the coding strand of the β-lactamase (Amp) gene, as shown in Fig. 2a. Fig. 2b also shows that poly(A) RNA from non-induced leukocytes, added to the hybridization at a similar level as induced poly(A) RNA, did not protect the labeled IF cDNA probe.

3) Nucleotide sequence analysis.

Hif-2h DNA was cleaved by an appropriate restriction enzyme, labeled with [^{32}P] phosphate at the 5' termini, and digested with a second restriction enzyme to yield fragments labeled at only one 5' end; the isolated fragments were sequenced by the Maxam-Gilbert procedure (Maxam and Gilbert, 1977). Fig. 1 shows the fragments analyzed in this fashion. Each stretch of the cDNA insert was sequenced from both strands, and each restriction site which served as labeled terminus was sequenced using a fragment spanning it. The nucleotide sequence thus obtained is shown in Fig. 3. The heteropolymeric part of the insert is flanked by 23 G residues at the 5' end and by 7 A residues (probably reflecting the poly(A) terminus of the mRNA) followed by 15 C residues at the 3' terminus. An AUG initiation triplet in position 57-59 and a UAA termination triplet at position 624-626 define a reading frame uninterrupted by nonsense codons. Both other reading frames contain 18 and 12 nonsense codons, respectively. The only other sequences flanked by an AUG

(or GUG) and by a termination triplet, which could code for a polypeptide of 25 amino acids or more, lie in different reading frames, between nucleotides 226 and 304, 640 and 778, and 683 and 743, respectively.

Hood and his colleagues have recently determined the sequence of 35 amino terminal amino acids of human lymphoblastoid interferon (Zoon et al., 1980; M. Hunkapiller and L. Hood, personal communication). In Fig. 5 the sequence of human lymphoblastoid IF (B) is aligned with the amino acid sequence determined by the major reading frame of the Hif-2h nucleotide sequence (A) such that the amino terminal amino acid of the former coincides with the amino acid coded for by the 24th codon of the latter. Extensive coincidence is found: 26 of 35 positions have identical amino acids. This confirms the assignment of the reading frame.

DISCUSSION

Cloned cDNA generated from poly(A) RNA by commonly used procedures (Efstratiadis et al., 1977) lacks 5' terminal nucleotides and may even contain artifactual sequences (Richards et al., 1979). It is therefore not certain whether the first AUG of the cloned human Le IF cDNA Hif-2h, which is located 57 nucleotides downstream from the 5' terminus of the heteropolymeric sequence, in fact corresponds to the first AUG on the mRNA.

Bearing these reservations in mind, we shall assume, until further experimental evidence becomes available, that this is the case.

In eukaryotic mRNAs the first AUG triplet from the 5' terminus is usually the initiation site for protein synthesis (Kozak, 1978). The codon in the cloned human Le IF cDNA correspondi to the first amino acid of lymphoblastoid interferon is 22 codons downstream from the first AUG (and 14 codons downstream from the second one) indicating that the sequence coding for interferon may be preceded by a sequence determining a signal peptide of 23 (or less likely 15) amino acids. The longer of the presumptive signal sequences contains an uninterrupted series of 11 hydrophobic amino acids (and the shorter one, one of 6). This accumulation of hydrophobic residues is characteristic of signal sequences (cf. Davis and Tai, 1980). The presumptive cleavage site between signal and interferon sequence lies between a Gly and a Cys residue. It is striking that in the case of E.coli prelipoprotein, cleavage occurs between the same two amino acids (Inouye et al., 1977). It will be interesting to determine whether the postulated preinterferon exists, and if so, whether it is correctly processed in E.coli, especially in view of our finding (S. Nagata, unpublished results) that about 50% of the interferon activity produced in E.coli can be released by osmotic shock and is therefore located in the periplasmic space (Anraku, 1968).

The sequence corresponding to (mature) Le IF polypeptide comprises 498 nucleotides, which code for 166 amino acids. Assuming that there is no carboxyterminal processing, the molecular weight of the interferon polypeptide, as calculated from Table 1, is 19'388. The base composition of the coding sequence is 50% GC; the codon usage within the interferon coding sequence (Table 2) is in reasonable agreement with that compiled for mammalian mRNAs in general (Grantham et al., 1980); the deviations observed may be ascribed to the small numbers involved.

The 3' non-coding region consists of 242 nucleotides; this length is intermediate between that of chicken ovalbumin mRNA (637 residues) (McReynolds et al., 1978) and rat insulin mRNA (53) (Ullrich et al., 1977). The high AT content (69%) is similar to that found for the corresponding segment of mouse β-globin minor mRNA (63%) (Konkel et al., 1979); the AT content of eukaryotic 3' non-coding regions range from 94% in mRNA yeast mitochondrial ATPase (Hensgens et al., 1979) to 42% in bovine ACTH-βLPH mRNA (Nakanishi et al., 1979). No striking homologies to 3' non-coding regions of other mRNAs were noted, except for the AATAAA(AC) sequence 18-27 nucleotides upstream from the poly(A) sequence, found previously (Proudfoot and Brownlee, 1976) in almost all eukaryotic mRNAs examined, at about the same relative position.

The comparison of the first 35 amino acids of lymphoblastoid interferon (Zoon et al., 1980; M. Hunkapiller and L. Hood, personal communication) and the sequence deduced from Hif-2h (Fig. 5) shows 9 differences. In all cases, the codons for the differing amino acids could be related by one-base changes. The amino acid compositions (Table 1) determined directly for lymphoblastoid interferon on the one hand and deduced from the Hif-2h sequence on the other, show striking differences in regard to their content of Gly, Pro, Cys and Met. These differences are too large to be explained by polymorphism; most likely we are dealing with the products of two non-allelic genes, since the degree of divergence of the two proteins (26% mismatch) is similar to that between, for example, human and sheep β globin (23% mismatch). We have recently surveyed our human leukocyte cDNA clone bank and identified a hybrid plasmid (2-pBR322(Pst)HcIF-II-206, or Hif-II206 for short) which also directs synthesis of interferon activity in E.coli and has a different restriction pattern than Hif-2h (M. Streuli and M. Schwarzstein, unpublished results). This clone represents a second leukocyte interferon gene (Le-IF II), differing from the one (Le-IF I) which corresponds to Hif-2h. The amino acid composition of an IF preparation from human leukocytes (Rubinstein et al., 1979) agrees somewhat better than that of lymphoblastoid IF with the amino acid composition deduced for Le-IF I (Table 1).

Taniguchi and his colleagues prepared cDNA from induced fibroblast poly(A) RNA and selected presumptive interferon cDNA clones by hybridization techniques (Taniguchi et al., 1979). The nucleotide sequence of one such clone was determined and could be correlated (Taniguchi et al., 1980) with the sequence of the 13 amino terminal amino acids of fibroblast

interferon (Knight et al., 1980). The striking structural homologies between the leukocyte and fibroblast interferon cDNA sequences will be analyzed elsewhere.

ACKNOWLEDGEMENTS

We thank A. van Ooyen for helpful suggestions concerning the sequencing, W. Boll and A. Kiss for restriction enzymes, J. Ecsödi for plasmid DNA and restriction analyses and M. Hunkapiller, L. Hood and T. Taniguchi for unpublished information. S.P. is a Scholar of the Leukemia Society of America. The work was supported by Biogen S.A. and the Schweizerische Nationalfonds.

REFERENCES

Anraku, Y., Transport of sugars and amino acids in bacteria, J. Biol. Chem. 243 (1968) 3116-3122.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.W., Construction and characterization of new cloning vehicles, II. A multipurpose cloning system, Gene, 2 (1977) 95-113.

Davis, B.D. and Tai, P.-C., The mechanism of protein secretion across membranes, Nature, 283 (1980) 433-438.

Efstratiadis, A., Kafatos, F.C. and Maniatis, T., The primary structure of rabbit β-globin mRNA as determined from cloned DNA, Cell, 10 (1977) 571-588.

Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pave, A., Codon catalog usage and the genome hypothesis, Nucl. Acids Res. 8 (1980) r49-r62.

Hensgens, L.A.M., Grivell, L.A., Borst, P. and Bos, J.L., Nucleotide sequence of the mitochondrial structural gene for subunit 9 of yeast ATPase complex, Proc. Natl. Acad. Sci. USA 76 (1979) 1663-1667.

Inouye, S., Wang, S., Sekizawa, J., Halegoua, S. and Inouye, M., Amino acid sequence for the peptide extension on the prolipoprotein of the Escherichia coli outer membrane, Proc. Natl. Acad. Sci. USA 74 (1977) 1004-1008.

Konkel, D.A., Maizel, J.V., jr, and Leder, P., The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes, Cell, 18 (1979) 865-873.

Kozak, M., How do eukaryotic ribosomes select initiation regions in messenger RNA?, Cell, 15 (1978) 1109-1123.

Loening, U.E., The fractionation of high-molecular weight ribonucleic acid by polyacrylamide gel electrophoresis, Biochem. J., 102 (1967) 251-257.

Maxam, A.M. and Gilbert, W., A new method for sequencing DNA, Proc. Natl. Acad. Sci. USA, 74 (1977) 560-564.

McReynolds, L., O'Malley, B.W., Nisbet, A.D., Fothergill, J.E., Givol, D., Fields, S., Robertson, M. and Brownlee, G.G., Sequence of chicken ovalbumin mRNA, Nature 273 (1978) 723-728.

Müller, W., Weber, H., Meyer, F. and Weissmann, C., Site-directed mutagenesis in DNA: Generation of point mutations in cloned β globin complementary DNA at the positions corresponding to amino acids 121 to 123.

Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. and Weissmann, C., Synthesis in E.coli of a polypeptide with human leukocyte interferon activity, Nature (1980) in press.

Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A.C.Y., Cohen, S.N. and Numa, S., Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor, Nature, 278 (1979) 423-427.

Proudfoot, N.J. and Brownlee, G.G., 3' non-coding region sequences in eukaryotic messenger RNA, Nature, 263 (1976) 211-214.

Richards, R.I., Shine, J., Ullrich, A., Wells, J.R.E. and Goodman, H.M., Molecular cloning and sequence analysis of adult chicken β globin cDNA, Nucl. Acids Res. 7 (1979) 1137-1146.

Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. and Pestka, S., Human leukocyte interferon: Production, purification to homogeneity, and initial characterization, Proc. Natl. Acad. Sci. USA 76 (1979) 640-644.

Smith, H.O. and Birnstiel, M.L., A simple method for DNA restriction site mapping, Nucl. Acids Res. 3 (1976) 2387-2398.

Sutcliffe, J.G., pBR322 restriction map derived from the DNA sequence, accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res., 5 (1978) 2721-2728.

Taniguchi, T., Sakai, M., Fujii-Kriyama, Y., Muramatsu, M., Kobayashi, S. and Sudo, T., Construction and identification of a bacterial plasmid containing the human fibroblast interferon gene sequence, Proc. Jap. Acad., 55 (1979) Ser.B, 464-469.

Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. and Muramatsu, M., The nucleotide sequence of human fibroblast interferon cDNA. Gene, in press (1980).

Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M. Rat insulin gene: construction of plasmids containing the coding sequences, Science, 196 (1977) 1313-1319.

Weaver, R.F. and Weissmann, C., Mapping of RNA by a modification of the Berk-Sharp procedure: the 5' termini of 15 S β-globin mRNA precursor and mature 10 S β-globin mRNA have identical map coordinates, Nucl. Acids Res., 7 (1979) 1175-1193.

Wiegand, R.C., Godson, G.N. and Radding, C.M., Specificity of the S_1 nuclease from Aspergillus oryzae, J. Biol. Chem., 250 (1975) 8848-8855.

Wilkie, N.M., Clements, J.B., Boll, W., Mantei, N., Londsdale, D. and Weissmann, C., Hybrid plasmids containing an active thymidine kinase gene of Herpes simplex virus 1, Nucl. Acids Res. 7 (1979) 859-877.

Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapiller, M.W. and Hood, L.E., Amino terminal sequence of the major component of human lymphoblastoid interferon, Science, 207 (1980) 527-528.

TABLE 1 Amino acid composition of leukocyte and lymphoblastoid interferon.[a)]

	Leukocyte IF, deduced from nucleotide sequence of Hif-2h cDNA[b)]	Lymphoblastoid IF[c)]	Leukocyt IF[d)]
AsN	6	15	15
Asp	11		
Thr	9	8	7 S
Ser	13	11	8
GlN	10	27	24
Glu	15		
Pro	6	11	6
Gly	3	11	5.5
Ala	10	11	8
Cys	5	2	3
Val	6	8	8
Met	6	3*	-
Ile	7	7	9
Leu	22	18	22
Tyr	4	4	5
Phe	8	7	9
His	3	4	3
Lys	8	10	12
Arg	12	10	7
Trp	2	1	1

TABLE 1 (cont.)

a) Values believed to differ significantly are underlined.

b) From Fig. 3.

c) From Zoon et al. (1980), except for the value marked with *, which was from L. Hood (personal communication).

d) From Rubinstein et al. (1979).

TABLE 2 Codon usage in the interferon coding sequence[a].

Codon		all MAM	IG	MAM-IG	Le IF
Arg	CGA	2	7	0	6
	CGC	9	4	12	0
	CGG	6	3	8	0
	CGU	7	1	9	0
	AGA	6	12	3	36
	AGG	11	9	11	30
Leu	CUA	8	11	7	6
	CUC	24	26	24	48
	CUG	51	16	68	42
	CUU	7	7	7	0
	UUA	4	11	1	12
	UUG	7	10	6	24
Ser	UCA	11	24	5	12
	UCC	20	17	22	24
	UCG	4	1	5	0
	UCU	17	29	11	24
	AGC	21	27	18	18
	AGU	17	33	9	0
Thr	ACA	13	26	7	18
	ACC	25	30	23	24
	ACG	7	4	9	0
	ACU	19	36	11	12

TABLE 2 (cont.)

Pro	CCA	12	20	7	6
	CCC	18	13	21	12
	CCG	8	5	9	0
	CCU	13	13	12	18
Ala	GCA	11	20	6	12
	GCC	34	22	40	18
	GCG	5	1	8	6
	GCU	23	21	24	24
Gly	GGA	10	20	5	12
	GGC	28	16	34	6
	GGG	11	9	11	0
	GGU	18	25	15	0
Val	GUA	3	5	1	0
	GUC	16	23	13	12
	GUG	35	17	44	18
	GUU	7	7	7	6
Lys	AAA	15	17	14	24
	AAG	45	24	56	24
Asn	AAC	29	29	29	24
	AAU	10	13	9	12
Gln	CAA	10	11	9	12
	CAG	32	30	32	48
His	CAC	22	8	29	6
	CAU	10	11	10	12

TABLE 2 (cont.)

Glu	GAA	22	21	23	36
	GAG	36	23	43	54
Asp	GAC	27	22	30	36
	GAU	18	23	16	30
Tyr	UAC	20	20	20	18
	UAU	15	17	14	6
Cys	UGC	12	6	16	6
	UGU	11	18	8	24
Phe	UUC	33	30	35	24
	UUU	16	13	18	24
Ile	AUA	4	7	3	0
	AUC	20	21	19	42
	AUU	11	18	8	0
Met	AUG	16	14	16	36
Trp	UGG	16	25	12	12

a) The values are expressed per 1000 amino acid residues. The data for "all MAM" (compiled from 18 mammalian sequences), "IG"(6 immunoglobulin sequences) and "MAM-1G" (12 non-immunoglobulin mammalian sequences) are from Grantham et al. (1980), those for human leukocyte interferon from Fig. 3.

FIGURE LEGENDS

Fig. 1 Strategy for the determination of the nucleotide sequence of Hif-2h DNA.

The restriction map was determined as outlined in the text and subsequently refined using the results of the nucleotide sequence analyes shown in Fig. 3. The filled circles represent labeled 5' termini, the solid arrows indicate the sequences read off the labeled fragments. The dashed lines represent regions not read off a particular fragment. Black box, interferon coding sequences; hatched box, putative signal sequence; white box, non-coding region. Straight lines, homopolymeric dG:dC flanking regions; wavy line, pBR322.

Fig. 2 Determination of the orientation of the Le IF coding sequence.

(a) An outline of the approach. Hif-2h DNA is cleaved asymmetrically within the IF cDNA sequence, at the BglII site. The 5' termini are labeled with [^{32}P] (filled circles) and the DNA cleaved with PstI. The labeled fragments, either separated or not, are denatured, hybridized with poly(A) RNA from IF-producing leukocytes, and the mixture digested with S_1 nuclease. If the coding sequence has the orientation

shown in the figure (as we found to be the case), the smaller (344 nucleotide) BglII fragment is recovered; if the orientation had been the opposite, the larger (578 nucleotide) BglII fragment would have been protected by the mRNA. +, sense strand; -, antisense strand; Amp, ampicillinase gene; Tet, tetracycline gene. The arrows indicate the direction of transcription. PN, polynucleotide kinase.

(b) Hif-2h DNA was cleaved, labeled and recleaved as outlined above. The specific ^{32}P-radioactivity was 1.3×10^6 cpm/pmol end. 0.015 pmol of probe in 5 µl hybridization buffer (80% formamide, 40 mM PIPES (pH 6.4), 0.4 M NaCl, 0.001 M EDTA) were denatured for 10 min at 60°C and transferred to a tube in which 0.15-5 µg Le poly(A) RNA had been dried down. In other similar experiments 25 µg oligo(C) were added to prevent protection of the dC residues of the probe by the dG residues of the probe or the DNA strand complementary to it. The mixture was heated 19 h at 48°C in a sealed capillary and transferred to 0.1 ml S_1 buffer (250 mM NaCl, 30 mM NaAc buffer (pH 4.5), 1 mM $ZnSO_4$) containing 1.5 µg denatured salmon sperm DNA and 55 units S_1 nuclease (prepared according to Wiegand et al. (1975) by A. Schamböck) were added. After 40 min at 30°C each sample was extracted with phenol-chloroform, 10 µg yeast

RNA were added and the nucleic acid was precipitated with 2 vol ethanol. The precipitate was dissolved in 5 µl 90% formamide, 2 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol, and analyzed by electrophoresis through a 5% polyacrylamide gel as described (Weaver and Weissmann, 1979). Autoradiography with an Ilford intensifying screen was for 18 h at -70°C. Lane 1, pBR322 cleaved with BspI and 5'-terminally labeled as marker (Sutcliffe, 1978); lane 2, the untreated, labeled probe; lanes 3-5, the labeled probe hybridized with 0.5, 1.5 and 5 µg uninduced poly(A) RNA, respectively lanes 6-8, the labeled probe hybridized to 0.5, 1.5 and 5 µg induced poly(A) RNA, respectively.

Fig. 3 The nucleotide sequence of Hif-2h IF cDNA.

The nucleotide sequence was determined as indicated in the Methods section. The amino acid sequence was deduced from the nucleotide sequence; lower case letters indicate the putative signal polypeptide. The MboI sites as well as the restriction targets marked with * were not cleavable.

Fig. 4 Autoradiogram of the sequence gel showing the presumed signal polypeptide coding region and the beginning of the interferon polypeptide.

Fragment b (Fig. 1) was degraded as described in the Methods section and analyzed on a 12% gel. The runs were for 2, 8, 18 and 26 h at 900 V. The four lanes for each run show, from left to right, degradations specific for G, A+G, C+T and C. Because of methylation at EcoRII sites, the Cs marked with asterisks are present as gaps in this sequence. Their existence was confirmed by analyses of the other strand.

Fig. 5 Comparison of the amino terminal amino acid sequence of lymphoblastoid interferon (determined experimentally) and leukocyte interferon (deduced from the Hif-2h cDNA nucleotide sequence).

The leukocyte interferon sequence (A) is from Fig. 3; the lymphoblastoid interferon sequence (B) is from Zoon et al. (1980), and M. Hunkapiller and L. Hood (personal communication). Dashes indicate identical amino acids.

Fig. 1

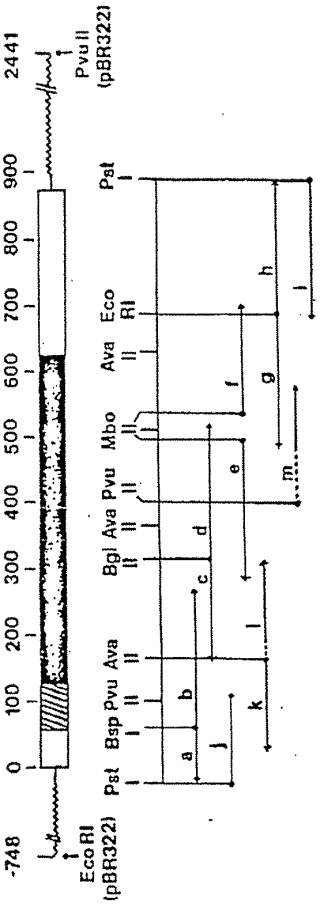
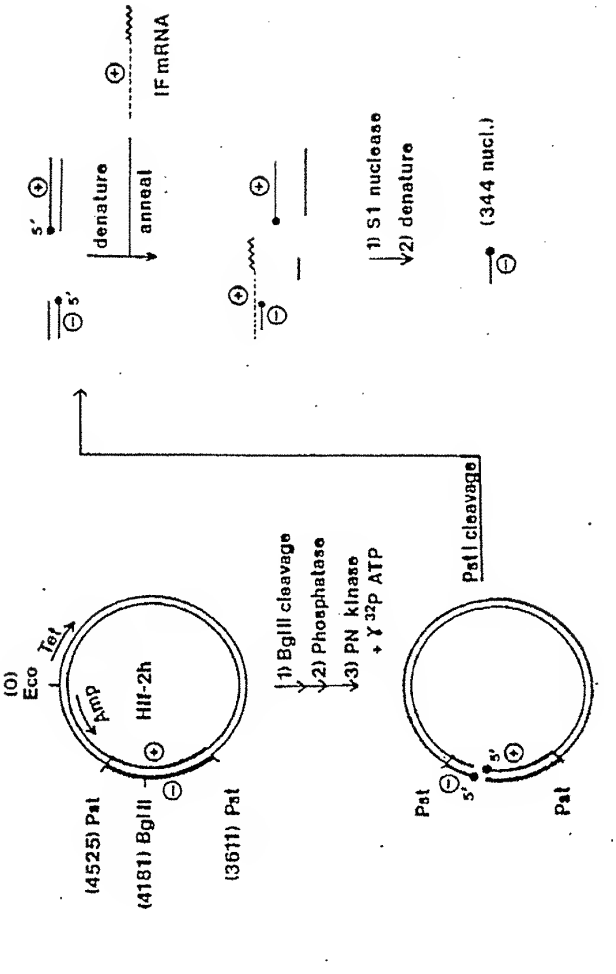
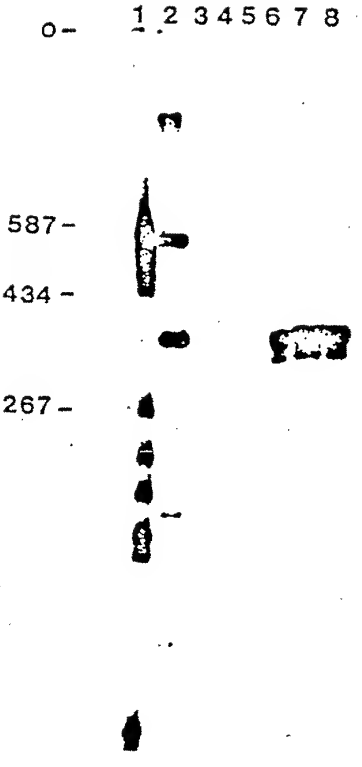

-748
0
100
200
300
400
500
600
700
800
900
2441
EcoRI (pBR322)
Pst
Bsp
Pvu II
Ava II
Bgl II
Ava II
Pvu II
Mbo II
Ava II
Eco RI
Pst
PvuII (pBR322)
a
b
c
d
e
f
g
h
i
j
k
l
m

Fig. 2a


(0)
Eco
Tet
pmP
Hif-2h
(4525) Pst
(4181) BglII
(3611) Pst
1) BglII cleavage
2) Phosphatase
3) PN kinase
+ γ 32P ATP
PstI cleavage
Pst
Pst
5'
denature
anneal
IF mRNA
1) S1 nuclease
2) denature
(344 nucl.)

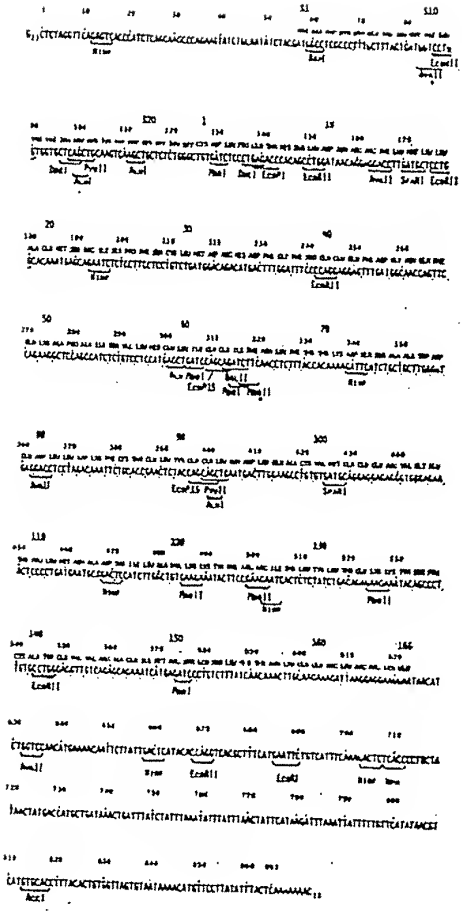
a

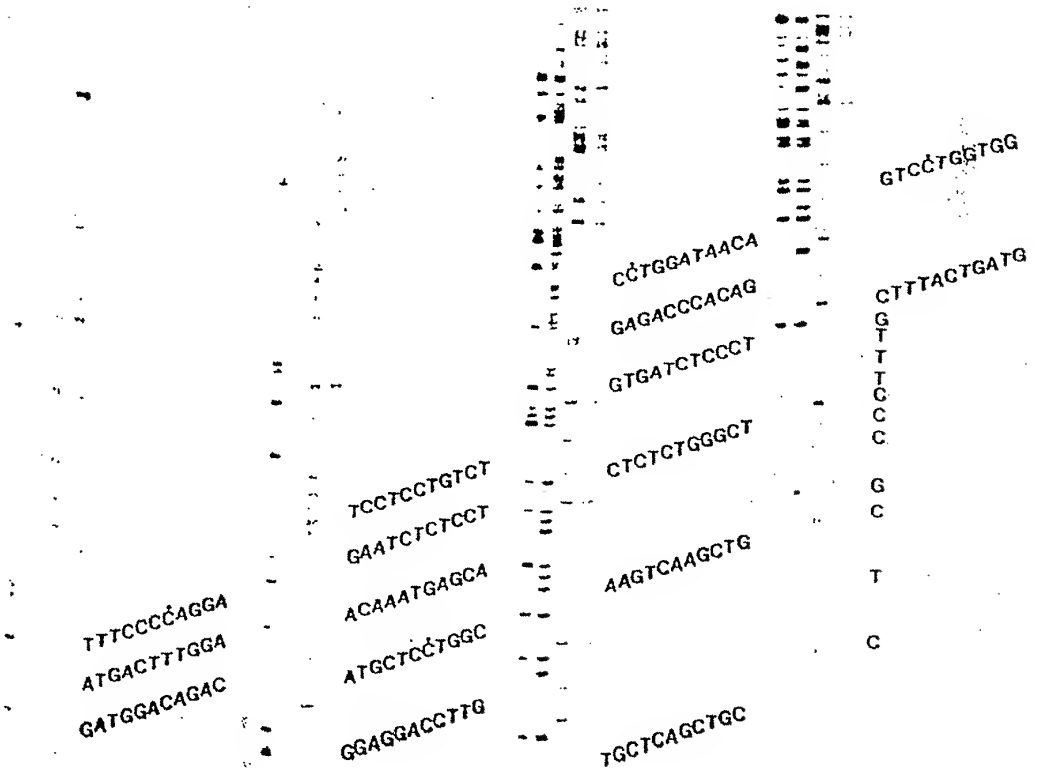
Fig. 2b



b

Fig. 3




TTTCCCCAGGA
ATGACTTTGGA
GATGGACAGAC
TCCTCCTGTCT
GAATCTCTCCT
ACAAATGAGCA
ATGCTCCTGGC
GGAGGACCTTG
CCTGGATAACA
GAGACCCACAG
GTGATCTCCCT
CTCTCTGGGCT
AAGTCAAGCTG
TGCTCAGCTGC
GTCCTGGTGG
CTTTACTGATG
G
T
T
C
C
C
G
C
T
C

A) Cys Asp Leu Pro Glu Thr His Ser Leu Asp Asn Arg Arg Thr Leu

B) Ser - - - Gln - - - - Gly - - - Ala -

A) Met Leu Leu Ala Gln Met Ser Arg Ile Ser Pro Ser Ser Cys Leu

B) Ile - - - - - Gly - - - Leu Phe - - -

A) Met Asp Arg His Asp

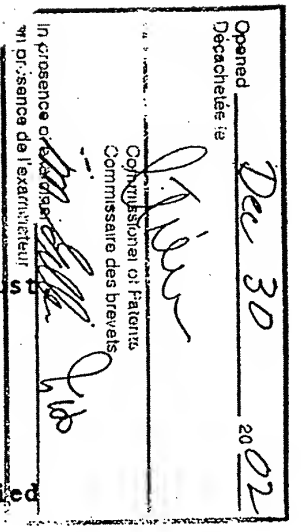
B) Lys - - - -

Fig. 5

TAB D

MAILING LIST for "Interferon" Manuscripts

Aebi, Bern
Aloni, Rehov
Anderson, Bethesda
Arber, Basel
Axel, Columbia
August, Bronx
Baltimore, Cambridge
Baron, Galveston
Bautz, Heidelberg
Bayev, Moscow
Berg, Palo Alto
Bernardi, Paris
Birnstiel, Zürich
Bishop, San Francisco
Biogen, Geneva
Borst, Amsterdam
Bové, Versailles
Brenner, Cambridge
Bres, Geneva
Bretscher, Cambridge
Brown, Baltimore
Burton, Oxford
Cantoni, Bethesda
Cantell, Helsinki
Carbon, Santa Barbara
Carter, Buffalo
Cawthorn, Geneva
Chamberlin, Berkeley
Chambon, Strasbourg
Chany, Paris (& Villejuif)
Clark, Cambridge
Clercq, Leuven
Coffin, Boston
Cohen, Stanford
Cohn, San Diego
Crick, San Diego
Cross, Wellcome
Curtis, Philadelphia
Curtiss III, Birmingham
Dahlberg, A.E., Bethesda
Dahlberg, J.E., Madison
Darnell, Bronx
Davies, Geneva
Davis, Stanford
de Maeyer, Orsay
Desmyter, Leuven
Doty, Cambridge
Duesberg, Berkeley
van der Eb, Leiden
Echols, Berkeley
Efstratiadis, Cambridge
Epstein, Geneva
Falcoff, Paris
Feix, Freiburg
Fiers, Gent
Finter, Beckenham
Flavell, London
Forget, Boston
Fresco, Princeton
Fraenkel-Conrat, Berkeley
Gallo, Bethesda
Garen, New Haven
Gesteland, Cold Spring Harbor
Gilbert, Cambridge
Gilham, Lafayette
Gilvarg, Princeton
Glaser, Berkeley
Goodman, San Francisco
Gresser, Villejuif
Gros, Paris
Gruber, Groningen
Grunberg-Manago, Paris
Hailey, Fish & Neane
Hanafusa, New York
Harris, Adelaide, S. Aust
Hartley, London
Haselkorn, Chicago
Havell, New York
Hindley, Bristol
Hinnen, Basel
Hogness, Boston
Hofman, Zürich
Hofschneider, Martinsried
Hohn, Basel
Holley, La Jolla
Hood, Pasadena
Horecker, Nutley
Hurwitz, Bronx
Jacob, Paris
Jeffreys, Leicester
Kaesberg, Madison
Kafatos, Cold Spring Harbor
Kamen, London
Kaziro, Tokyo
Kellenberger, Basel
Kerr, London
Khorana, Cambridge
Klingmüller, München
Knight, Wilmington
Konigsberg, New Haven
Koprowski, Philadelphia
Kornberg, Palo Alto
Kourilsky, Paris
Krakow, New York
Lawrason, Kenilworth
Leder, Bethesda
Lederberg, Stanford
Leibowitz, Bloomfield
Lengyel, New Haven
Levy, Bethesda
Libonati, Napoli
Lindenmann, Zürich
Lingrel, Cincinnati
Littauer, Rehovoth
Lodish, Cambridge
Luria, Cambridge
Lüscher, Bern
Mach, Geneva
Maniatis, Pasadena
Marks, New York
Marmur, Bronx
Martin, Bethesda
Merigan, Stanford
Murray, Heidelberg
Montagnier, Paris
Nathans, Baltimore



This is EXHIBIT FIERS-39
to
the Affidavit of Walter C. Fiers
sworn before me
this 19th day of November, 2001

Commissioner for Oath or Notary Public

Sugano EXHIBIT 2024
Sugano v. Goeddel
Interference No. 105,334 and 105,337

MAILING LIST for "Interferon" Manuscripts

Aebi, Bern
Aloni, Rehov.
Anderson, Bethesda
Arber, Basel
Axel, Columbia
August, Bronx
Baltimore, Cambridge
Baron, Galveston
Bautz, Heidelberg
Bayev, Moscow
Berg, Palo Alto
Bernardi, Paris
Birnstiel, Zürich
Bishop, San Francisco
Biogen, Geneva
Borst, Amsterdam
Bové, Versailles
Brenner, Cambridge
Bres, Geneva
Bretscher, Cambridge
Brown, Baltimore
Burton, Oxford
Cantoni, Bethesda
Cantell, Helsinki
Carbon, Santa Barbara
Carter, Buffalo
Cawthorn, Geneva
Chamberlin, Berkeley
Chambon, Strasbourg
Chany, Paris (& Villejuif)
Clark, Cambridge
Clercq, Leuven
Coffin, Boston
Cohen, Stanford
Cohn, San Diego
Crick, San Diego
Cross, Wellcome
Curtis, Philadelphia
Curtiss III, Birmingham
Dahlberg, A.E., Bethesda
Dahlberg, J.E., Madison
Darnell, Bronx
Davies, Geneva
Davis, Stanford
de Maeyer, Orsay
Desmyter, Leuven
Doty, Cambridge
Duesberg, Berkeley
van der Eb, Leiden
Echols, Berkeley
Efstratiadis, Cambridge
Epstein, Geneva
Falcoff, Paris
Feix, Freiburg
Fiers, Gent
Finter, Beckenham
Flavell, London
Forget, Boston
Fresco, Princeton
Fraenkel-Conrat, Berkeley
Gallo, Bethesda
Garen, New Haven
Gesteland, Cold Spring Harbor

Gilbert, Cambridge
Gilham, Lafayette
Gilvarg, Princeton
Glaser, Berkeley
Goodman, San Francisco
Gresser, Villejuif
Gros, Paris
Gruber, Groningen
Grunberg-Manago, Paris
Hailey, Fish & Neane
Hanafusa, New York
Harris, Adelaide, S. Aust.
Hartley, London
Haselkorn, Chicago
Havell, New York
Hindley, Bristol
Hinnen, Basel
Hogness, Boston
Hofman, Zürich
Hofschneider, Martinsried
Hohn, Basel
Holley, La Jolla
Hood, Pasadena
Horecker, Nutley
Hurwitz, Bronx
Jacob, Paris
Jeffreys, Leicester
Kaesberg, Madison
Kafatos, Cold Spring Harbor
Kamen, London
Kaziro, Tokyo
Kellenberger, Basel
Kerr, London
Khorana, Cambridge
Klingmöller, München
Knight, Wilmington
Konigsberg, New Haven
Koprowski, Philadelphia
Kornberg, Palo Alto
Kourilsky, Paris
Krakow, New York
Lawrason, Kenilworth
Leder, Bethesda
Lederberg, Stanford
Leibowitz, Bloomfield
Lengyel, New Haven
Levy, Bethesda
Libonati, Napoli
Lindenmann, Zürich
Lingrel, Cincinnati
Littauer, Rehovoth
Lodish, Cambridge
Luria, Cambridge
Lüscher, Bern
Mach, Geneva
Maniatis, Pasadena
Marks, New York
Marmur, Bronx
Martin, Bethesda
Merigan, Stanford
Murray, Heidelberg
Montagnier, Paris
Nathans, Baltimore

Opened
Décacheté le
Dec 30 2002
Commissioner of Patents
Commissaire des brevets
In presence of examiner
En présence de l'examinateur

This is EXHIBIT FIERS-39
to
the Affidavit of Walter C. Fiers
sworn before me
this 29th day of November, 2001

Commissioner for Oath or Notary Public

Mailing list for "Interferon" manuscripts - page 2

Ochoa, Nut
Oeschger, Baltimore
O'Malley, Houston
Parsons, Charlottesville
Paul, Glasgow
Pitha, Baltimore
Pestka, Nutley
Ptashne, Cambridge
Rajbhandari, Madison
Rauscher, Bethesda
Reich, New York
Revel, Rehovoth
Rich, Cambridge
Ross, Madison
Rowe, Bethesda
Ruddle, New Haven
Sanger, Cambridge
Schaller, Heidelberg
Schwartz, New York
Shapiro, Bronx
Sharp, Cambridge
Singer, Bethesda
Skalka, Nutley
Sinsheimer, Pasadena
Smith, Cambridge
Spahr, Geneva
Speyer, Storrs
Spiegelman, New York
Sgaramella, Pavia
Staehelin, Basel
Steitz, New Haven
Stent, Berkeley
Stewart, New York
Strander, Stockholm
Streisinger, Eugene
Stutz, Neuchâtel
Subak-Sharpe, Glasgow
Swetley, Vienna
Szybalski, Madison
Taniguchi, Tokyo
Temin, Madison
Tener, Vancouver
Thomas, Boston
Tissières, Geneva
Tocchini-Valentini, Napoli
Tonegawa, Basel
Tooze, Heidelberg
Veno, Düsseldorf
Vilček, New York
Viñuela, Madrid
Vogt, Los Angeles
Walker, Edinburgh
Warner, Irvine
Watson, Cold Spring Harbor
Weber, R., Bern
Weissman, New Haven
Williamson, London
Wittmann, Berlin
Yanofsky, Stanford
Zachau, München
Zillig, München
Zinder, New York

TAB E

SOUTH AUSTRALIAN INSTITUTE
OF TECHNOLOGY

TERRACE, ADELAIDE, SOUTH AUSTRALIA · · · · · · · TELEPHONE 223 3000

School of Pharmacy

10th April, 1980

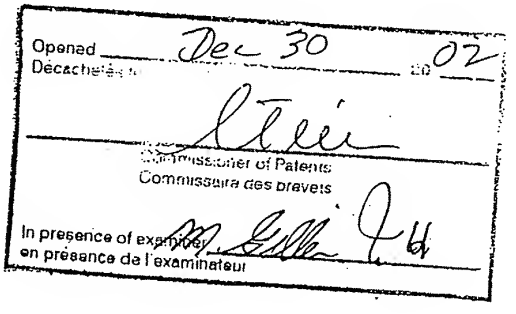
Dr. C. Weissmann,
Institut fur Molecularbiologie I,
Universitat Zurich,
HONGGERBERG. Zurich. 8093
Switzerland.

18. April 1980

Dear Dr. Weissmann,

Thank you again for the preprints. Fantastic work. Have you thought about the possible advantages of interferon production in B. subtilis? The leader peptides may be adequate for transport of the interferons out to the medium thus simplifying the purification of interferon (however proteases may be a problem). A sequence comparison realignment of sections of the interferons is enclosed. I was quite surprised by the similarity in amino acid composition of bovine growth hormone and Le-IFI. However I could find little sequence homology.

I wish you all the very best in your work.



Yours sincerely,

Roy

R.J. Harris
Lecturer in Biochemistry

PS. Do you mind if I distribute the preprints within Australia? I will assume this is ok unless you notify me otherwise. I will be presenting data on applications of genetic engineering to a meeting in May 1980. I would love to work in your lab but this would perhaps be incompatible with my aim to produce interferon within Australia.
Would you be interested in attending the International Biochem. Conference in Perth (Western Australia) 1982?

This is EXHIBIT FIERS-42
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary

SOUTH AUSTRALIAN INSTITUTE
OF TECHNOLOGY

TERRACE, ADELAIDE, SOUTH AUSTRALIA • • • • • • • TELEPHONE 223 3000

School of Pharmacy

10th April, 1980

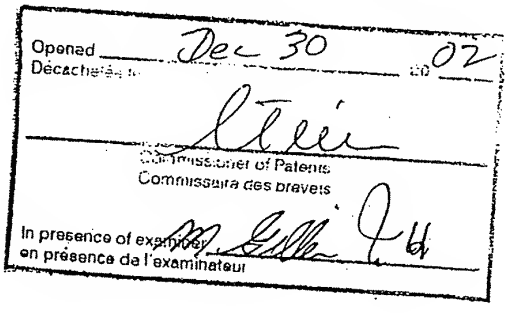
Dr. C. Weissmann,
Institut fur Molecularbiologie I,
Universitat Zurich,
HONGGERBERG. Zurich. 8093
Switzerland.

18. April 1980

Dear Dr. Weissmann,

Thank you again for the preprints. Fantastic work. Have you thought about the possible advantages of interferon production in B. subtilis? The leader peptides may be adequate for transport of the interferons out to the medium thus simplifying the purification of Interferon (however proteases may be a problem). A sequence comparison realignment of sections of the interferons is enclosed. I was quite surprised by the similarity in amino acid composition of bovine growth hormone and Le-IFI. However I could find little sequence homology.

I wish you all the very best in your work.



Yours sincerely,

Roy

R.J. Harris
Lecturer in Biochemistry

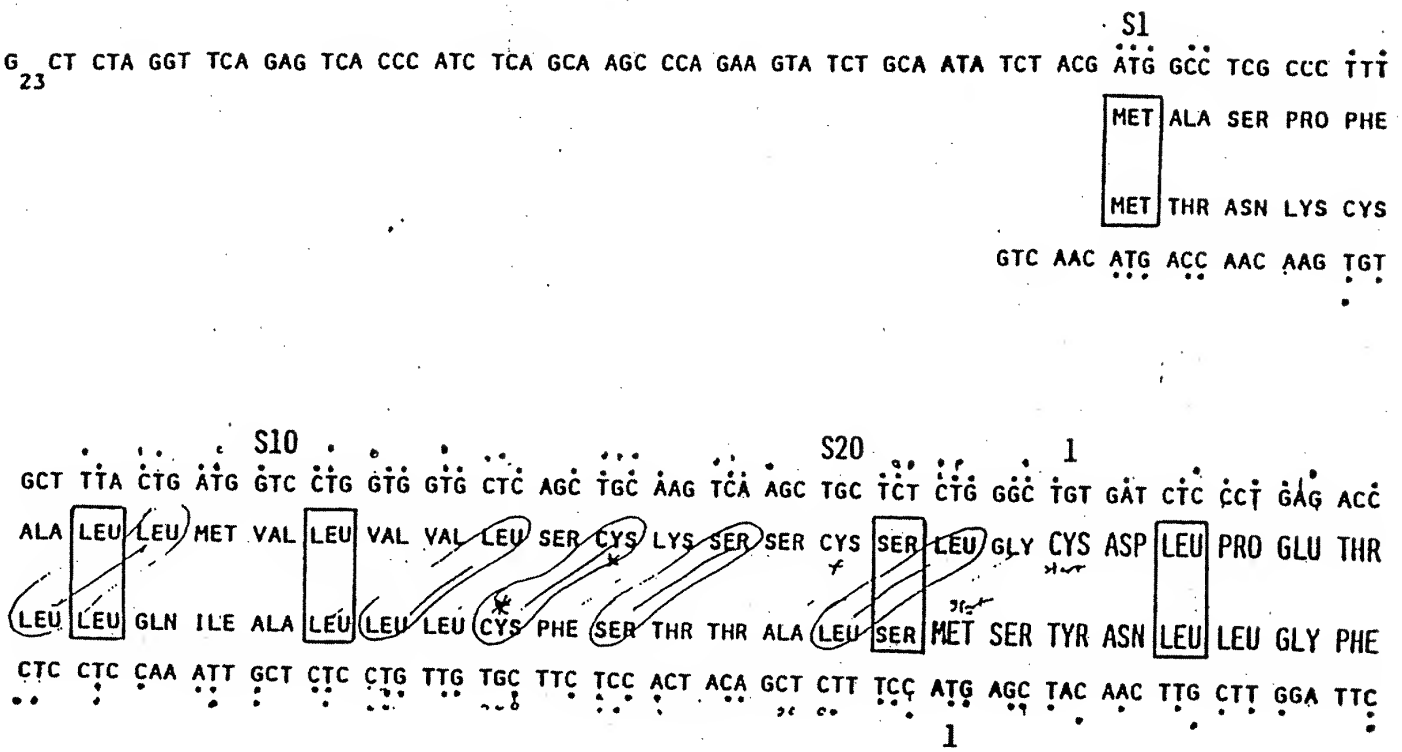
PS. Do you mind if I distribute the preprints within Australia? I will assume this is ok unless you notify me otherwise. I will be presenting data on applications of genetic engineering to a meeting in May 1980. I would love to work in your lab but this would perhaps be incompatible with my aim to produce interferon within Australia.
Would you be interested in attending the International Biochem. Conference in Perth (Western Australia) 1982?

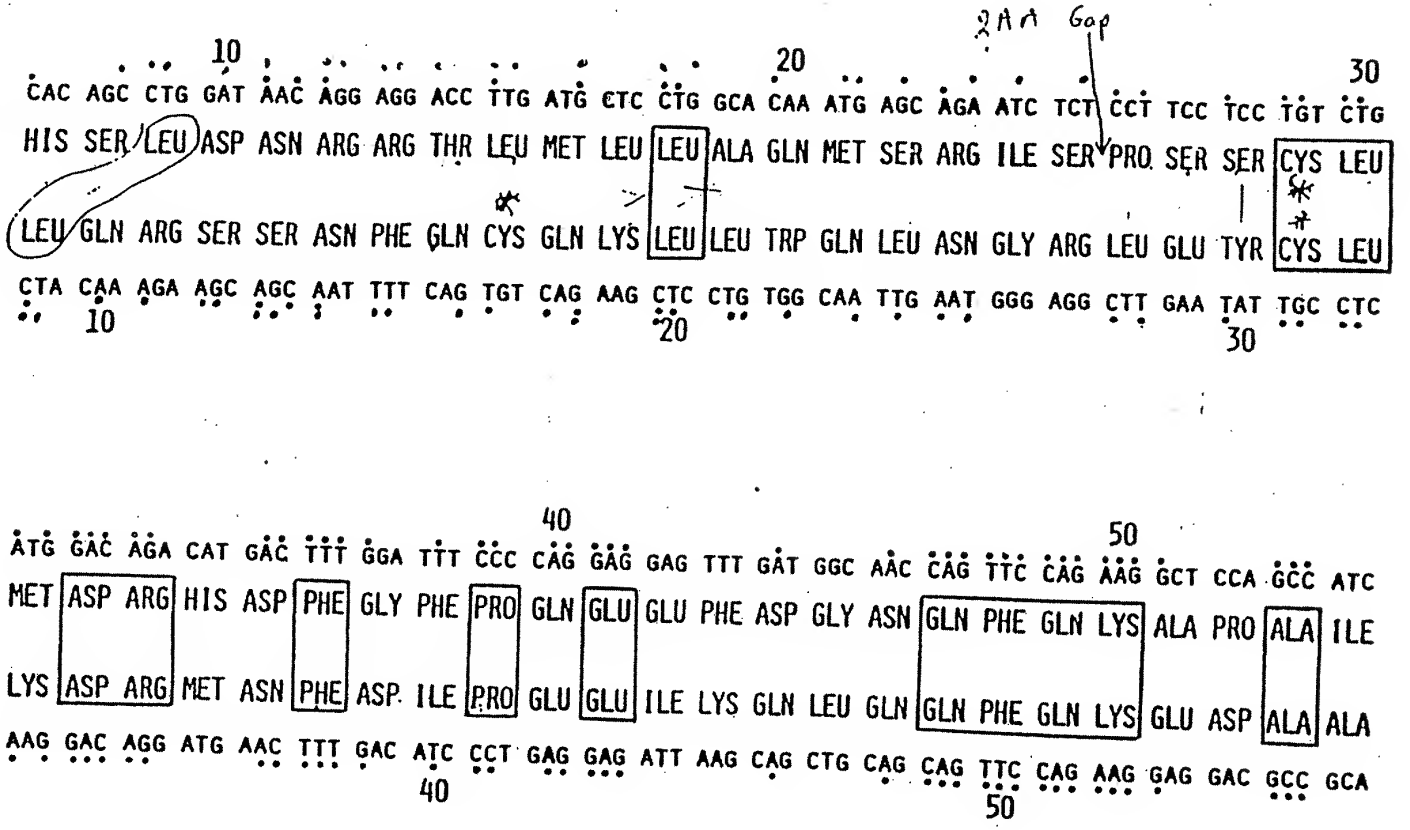
This is EXHIBIT FIERS-42
to
the Affidavit of Walter C. Fiers
sworn before me
this 19 th day of November, 2001

Commissioner for Oath or Notary Public

Fig. 1

do compiler run




2AA Gap
10
20
30
CAC AGC CTG GAT AAC AGG AGG ACC TTG ATG CTC CTG GCA CAA ATG AGC AGA ATC TCT CCT TCC TCC TGT CTG
HIS SER LEU ASP ASN ARG ARG THR LEU MET LEU LEU ALA GLN MET SER ARG ILE SER PRO SER SER CYS LEU
LEU GLN ARG SER SER ASN PHE GLN CYS GLN LYS LEU LEU TRP GLN LEU ASN GLY ARG LEU GLU TYR CYS LEU
CTA CAA AGA AGC AGC AAT TTT CAG TGT CAG AAG CTC CTG TGG CAA TTG AAT GGG AGG CTT GAA TAT TGC CTC
10
20
30
40
50
ATG GAC AGA CAT GAC TTT GGA TTT CCC CAG GAG GAG TTT GAT GGC AAC CAG TTC CAG AAG GCT CCA GCC ATC
MET ASP ARG HIS ASP PHE GLY PHE PRO GLN GLU GLU PHE ASP GLY ASN GLN PHE GLN LYS ALA PRO ALA ILE
LYS ASP ARG MET ASN PHE ASP ILE PRO GLU GLU ILE LYS GLN LEU GLN GLN PHE GLN LYS GLU ASP ALA ALA
AAG GAC AGG ATG AAC TTT GAC ATC CCT GAG GAG ATT AAG CAG CTG CAG CAG TTC CAG AAG GAG GAC GCC GCA
40
50

TAB F

Gene, 10 (1980) 1–10

The nucleotide sequence of a cloned human leukocyte interferon cDNA

(Recombinant DNA; pBR322 plasmid vector; restriction endonucleases; preinterferon amino acid sequence)

Ned Mantei, Marco Schwarzstein, Michel Streuli, Sandra Panem *, Shigekazu Nagata and Charles Weissmann

Institut für Molekularbiologie I, Universität Zürich, Hönggerberg, 8093 Zürich (Switzerland)

(Received March 12th, 1980)
(Accepted March 20th, 1980)

SUMMARY

We have determined the nucleotide sequence of the human leukocyte interferon cDNA carried in hybrid plasmid Z-pBR322(*Pst*)/HcIF-2h, which has been shown to direct the formation of a polypeptide with human leukocyte interferon activity (Nagata et al., 1980). The 910 base pair insert contains a 567 (or 543) base pair coding sequence, which determines a putative preinterferon polypeptide consisting of a signal peptide of 23 (or less likely 15) amino acids, followed by an interferon polypeptide of 166 amino acids (calculated molecular weight, 19 390). The coding sequence is preceded by a (most likely incomplete) 56 bp leader and followed by a 242 bp trailer and seven A residues from the poly(A) tail: A comparison of the sequence of 35 amino terminal amino acids of lymphoblastoid interferon (Zoon et al., 1980; M. Hunkapiller and L. Hood, personal communication) and the corresponding sequence deduced for leukocyte interferon revealed 9 differences. This suggests that these two interferons are encoded by two non-allelic genes.

INTRODUCTION

We have recently described the isolation of a hybrid plasmid Z-pBR322(*Pst*)/HcIF-2h, or Hif-2h for short, which contains a cDNA sequence coding for human leukocyte interferon. The hybrid DNA was identified by its capacity to (a) hybridize with human leukocyte interferon mRNA and (b) to direct the synthesis in *E. coli* of a protein with properties of human leukocyte interferon (Nagata et al., 1980).

In this paper we report the nucleotide sequence of the 910 bp insert of Hif-2h. Two AUG triplets and a UAA termination codon, all in the same reading frame, define a stretch of 567 or 543 nucleotides which encodes a polypeptide of 166 amino acids corresponding to the interferon polypeptide proper, preceded by 23 or 15 amino acids, which may constitute a signal sequence. The coding region is flanked at the 5′ end by 79 nucleotides, 23 of which are terminal G residues, and at the 3′ end by 264 nucleotides, 15 of which are terminal C residues.

MATERIALS AND METHODS

(a) Plasmids, enzymes and ATP

Plasmid DNA was prepared by method B described in Wilkie et al. (1979). *Eco*RI was a gift from W. Boll and *Bsp*I from A. Kiss. All other restriction enzymes

* Present address: Department of Pathology, University of Chicago, Chicago, IL 60637 (U.S.A.).

Abbreviations: bp, base pairs; LeIF, leukocyte interferon.

were purchased from New England Biolabs and used in essence as recommended by the supplier (except that 200 μg/ml gelatin replaced bovine serum albumin in the buffers).

(b) 5′-terminal labeling of DNA

Restricted DNA (20 μg) was extracted with phenol, precipitated with ethanol, dissolved in 0.05 M Tris · HCl (pH 8), and passed over a small column of Chelex-100. Fragments with flush or 5′-overhanging ends were dephosphorylated by treatment with 0.2 units calf intestinal alkaline phosphatase (Boehringer) per pmol DNA 5′-ends in 200 μl 0.05 M Tris · HCl (pH 8) for 60 min at 37°C. The enzyme was inactivated by heating 60 min at 65°C. For DNA fragments with 3′-overhanging ends, bacterial alkaline phosphatase (Worthington) was used as described (Maxam and Gilbert, 1977), except that incubation was at 65°C for 30 min. The dephosphorylated DNA was purified by adsorption to and elution from DEAE-cellulose as described (Müller et al., 1978) or subjected to polyacrylamide gel electrophoresis where required (see below). Fragments recovered from a polyacrylamide (or agarose) gel in 0.15 M NaCl, 0.05 M Tris · HCl (pH 8) were adsorbed to a 0.1-ml hydroxyapatite (Biorad HTP) column, washed 4 times with 1 ml of 0.1 M potassium phosphate buffer (pH 7) and eluted with 0.3 ml 1 M potassium phosphate buffer (pH 7). The solution was diluted tenfold and the DNA adsorbed to DEAE cellulose and recovered as described (Müller et al., 1978).

After ethanol precipitation, the DNA was 5′-terminally labeled with [γ-^{32}P]ATP (carrier-free, prepared by unpublished procedure of I. Kennedy; 12–34 μCi/pmol DNA end) and polynucleotide kinase (New England Biolabs or P-L Biochemicals Inc.) essentially as described (Maxam and Gilbert, 1977), except that the DNA was not denatured before the kinase reaction. Specific activities of 1–1.5 μCi [^{32}P]phosphate/mol DNA 5′-ends were obtained.

(c) Nucleotide sequence determination

For sequencing, labeled fragments were cleaved with a second restriction enzyme and the products separated by electrophoresis through a 5% polyacrylamide gel in Tris-borate-EDTA buffer. The desired fragments were extracted from the gel and purified as described (Müller et al., 1978). The various fragments for sequencing were prepared as follows (the number indicates the nominal fragment chain length in bp, the ends are indicated by enzyme cuts, the labeled site is indicated by an asterisk, and the letters in parentheses refer to the arrows shown in Fig. 1). (a) and (b), cleavage of Hif-2h with *Bsp*I, isolation by 5% polyacrylamide gel electrophoresis in Loening's buffer (Loening, 1967) of *Bsp-Bsp*-232 (for (a)) and *Bsp-Bsp*-949 (for (b)), labeling, cleavage with *Pst*I, isolation of (a) *Bsp*-Pst*-83 and (b) *Bsp*-Pst*-827. (c) and (d), cleavage of Hif-2h with *Bgl*II, labeling, cleavage with *Pst*I, isolation of (c) *Bgl*-Pst*-336 and (d) *Bgl*-Pst*-570. (e) and (f), cleavage of Hif-2h with *Mbo*II, labeling, digestion with *Pst*I and *Hin*dII (to cleave an interfering 350 bp pBR322 fragment), isolation of (e) *Mbo*-Pst*-519 and (f) *Mbo*-Pst*-351. (g) and (h), cleavage of Hif-2h with *Eco*RI, labeling, cleavage with *Pst*I, isolation of (g) *Eco*-Pst*-708 and (h) *Eco*-Pst*-198. (i) and (j), cleavage of Hif-2h with *Pst*I, labeling, cleavage with *Bgl*II, isolation of (i) *Pst*-Bgl*-570 and (j) *Pst*-Bgl*-336. (k) and (l), cleavage of Hif-2h with *Ava*II, labeling, cleavage with *Pst*I and *Bgl*II, isolation of (k) *Ava*-Pst*-186 and (l) *Ava*-Bgl*-147. (m) cleavage of plasmid with *Pvu*II, labeling, cleavage with *Pst*I and *Bgl*II, isolation of *Pvu*-Pst*-486. The fragments were degraded according to Maxam and Gilbert (1977), with the modifications described in protocols provided by the same authors in September, 1978. The products were fractionated on 0.1 × 25 × 36 cm 12% polyacrylamide gels (acrylamide/bisacrylamide = 18/1) in 50 mM Tris-borate, 1 mM EDTA (pH 8.3), with runs of 2, 8, 18 and 26 h at 900 V following a 6 h prerun at 700 V. Best results were obtained when the gels were kept at room temperature 2–3 days before use.

RESULTS

(1) Physical map of Hif-2h DNA

Hif-2h consists of dC-elongated human Le IF cDNA joined to pBR322 (Bolivar et al., 1977) which had been cleaved with *Pst*I and elongated with dG residues. A physical map was prepared by measuring the lengths of the fragments generated by single cleavages with *Eco*RI, *Bsp*I, *Pst*I and *Mbo*II and

double cleavages with *Pst*I on the one hand and *Eco*RI, *Bgl*II, *Bsp*I and *Mbo*II on the other, as well as with *Eco*RI and *Bgl*II, and *Eco*RI and *Mbo*II. In addition, DNA fragments which were ^{32}P-labeled at one 5′-end, were partially digested with a variety of restriction enzymes, and the lengths of the labeled products determined (Smith and Birnstiel, 1976). The resulting preliminary map was used as a basis for the nucleotide sequence analysis; the map shown in Fig. 1 was refined (see also Fig. 3) using the results of the nucleotide sequence analysis described below. No restriction targets for *Bgl*I, *Kpn*I, *Hae*II, *Xho*I, *Pvu*I, *Xba*I, *Pst*I, *Bst*EII, *Bam*HI, *Hin*dII, *Sal*I, *Hin*dIII, *Hpa*II, *Taq*I, *Hga*I, *Tha*I (*Tac*I), *Hpa*I or *Hha*I were found in the insert; there were single sites for *Bsp*I, *Bgl*II and *Eco*RI, and two sites for *Pvu*II. One each of 4 *Ava*II and 4 *Mbo*II targets (marked with * in Fig. 3) was not cleavable by the cognate enzyme, perhaps because of methylated bases in adjacent *Eco*RII and *Mbo*I sites. The *Mbo*I sites were not cleavable.

(2) The orientation of the coding sequence

In order to determine the orientation of the coding strand relative to pBR322 the experiment outlined in Fig. 2a was carried out. The hybrid plasmid Hif-2h was cleaved at the single *Bgl*II site, 5′-terminally labeled with [^{32}P]phosphate and digested with *Pst*I to yield 336/344 and 578/570 bp radioactive fragments. The fragments were denatured, annealed with poly(A) RNA from induced leukocytes, and the mixture was treated with S_1 nuclease. The resulting products were denatured and analyzed by polyacrylamide gel electrophoresis. As shown in Fig. 2b, a ^{32}P-labeled fragment of about 340 nucleotides was protected. In another, similar experiment the labeled fragments were first separated and then annealed individually with poly(A) RNA: the shorter, but not the longer probe was protected against S_1 nuclease (data not shown). These experiments identify the 5′-labeled 344 nucleotide strand as the minus strand, i.e. the strand complementary to the mRNA. Therefore, the orientation of the insert is such that the coding strand of the IF cDNA is a continuation of the coding strand of the β-lactamase (Amp) gene, as shown in Fig. 2a. Fig. 2b also shows that poly(A) RNA from non-induced leukocytes, added to the hybridization at a similar level as induced poly(A) RNA, did not protect the labeled IF cDNA probe.

(3) Nucleotide sequence analysis

Hif-2h DNA was cleaved by an appropriate restriction enzyme, labeled with [^{32}P]phosphate at the 5′ termini, and digested with a second restriction enzyme to yield fragments labeled at only one 5′ end; the isolated fragments were sequenced by the procedure of Maxam and Gilbert (1977). Fig. 1 shows the fragments analyzed in this fashion. Each stretch of the cDNA insert was sequenced from both strands,

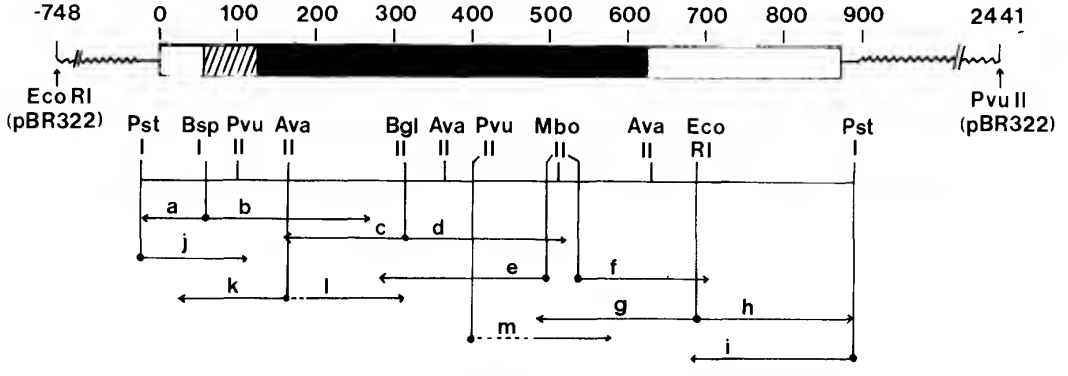


Fig. 1. Strategy for the determination of the nucleotide sequence of Hif-2h DNA. The restriction map was determined as outlined in the text and subsequently refined using the results of the nucleotide sequence analyses shown in Fig. 3. The filled circles represent labeled 5′ termini, the solid arrows indicate the sequences read off the labeled fragments. The dashed lines represent regions not read off a particular fragment. Upper map: numbers indicate bp; black segment, interferon coding sequences; hatched segment, putative signal sequence; white segment, non-coding region; straight lines, next to rectangle, homopolymeric dG:dC flanking regions: wave lines, pBR322.

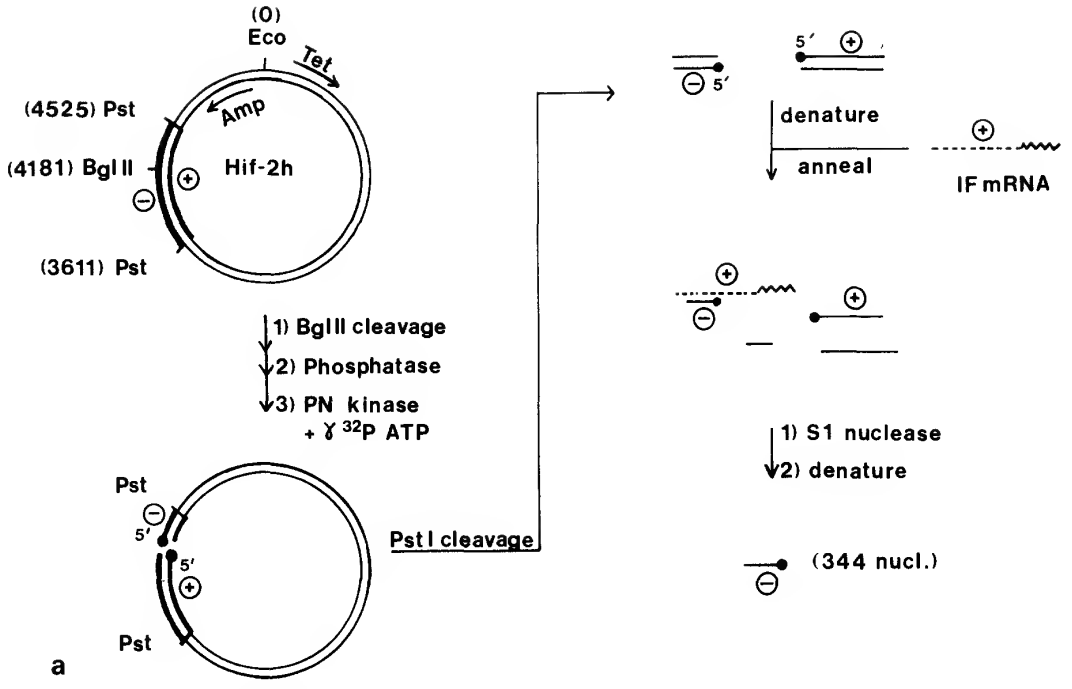


Fig. 2. Determination of the orientation of the LeIF coding sequence. (a) An outline of the approach. Hif-2h DNA is cleaved asymmetrically within the LeIF cDNA sequence, at the *Bgl*II site. The 5′ termini are labeled with ^{32}P (filled circles) and the DNA cleaved with *Pst*I. The labcled fragments, either separated or not, are denatured, hybridized with poly(A) RNA from LeIF-producing leukocytes, and the mixture digested with S_1 nuclease. If the coding sequence has the orientation shown in the figure (as we found to be the case), the smaller (344 nucleotide) *Bgl*II fragment is recovered; if the orientation had been the opposite, the larger (578 nucleotide) *Bgl*II fragment would have been protected by the mRNA. +, sense strand; −, antisense strand; Amp, ampicillinase genc; Tet, tetracycline-resistance gene. The arrows indicate the direction of transcription. PN, polynucleotide kinase. (b) Hif-2h DNA was cleaved, labeled and recleaved as outlined in (a). The specific ^{32}P-radioactivity was 1.3×10^6 cpm/pmol of end. 0.015 pmol of probe in 5 μl hybridization buffer (80% formamide, 40 mM PIPES (pH 6.4), 0.4 M NaCl, 0.001 M EDTA) were denatured for 10 min at 60°C and transferred to a tube in which 0.15–5 μg Le poly(A) RNA had been dried down. In other similar experiments 25 μg oligo(C) were added to prevent protection of the dC residues of the probe by the dG residues of the probe or the DNA strand complementary to it. The mixture was heated 19 h at 48°C in a sealed capillary and transferred to 0.1 ml S_1 buffer (250 mM NaCl, 30 mM NaAc buffer (pH 4.5), 1 mM $ZnSO_4$) containing 1.5 μg denatured salmon sperm DNA and 55 units S_1 nuclease (prepared according to Wiegand et al. (1975) by A. Schamböck) were added. After 40 min at 30°C each sample was extracted with phenol-chloroform, 10 μg yeast RNA were added and the nucleic acid was precipitated with 2 vol. ethanol. The precipitate was dissolved in 5 μl 90% formamide, 2 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol, and analyzed by electrophoresis through a 5% polyacrylamide gel as described (Weaver and Weissmann, 1979). Autoradiography with an Ilford intensifying screen was for 18 h at −70°C. Lane 1, pBR322 cleaved with *Bsp*I and 5′-terminally labeled to serve as markers (Sutcliffe, 1978); lane 2, the untreated, labeled probe; lanes 3–5, the labeled probe hybridized with 0.5, 1.5 and 5 μg uninduced poly(A) RNA, respectively; lanes 6–8, the labeled probe hybridized to 0.5, 1.5 and 5 μg induced poly(A) RNA, respectively.

and each restriction site which served as labeled terminus was sequenced using a fragment spanning it. The nucleotide sequence thus obtained is shown in Fig. 3. The heteropolymeric part of the insert is flanked by 23 G residues at the 5′-end and by 7 A residues (probably reflecting the poly(A) terminus of the mRNA) followed by 15 C residues at the 3′ terminus. An AUG initiation triplet in position 57–59 and a UAA termination triplet at position 624–626 define a reading frame uninterrupted by nonsense codons. Both other reading frames contain 18 and 12 nonsense codons, respectively. The only other sequences flanked by an AUG (or GUG) and by a termination triplet, which could code for a polypeptide of 25 amino acids or more, lie in different reading frames, between nucleotides 226 and 304,

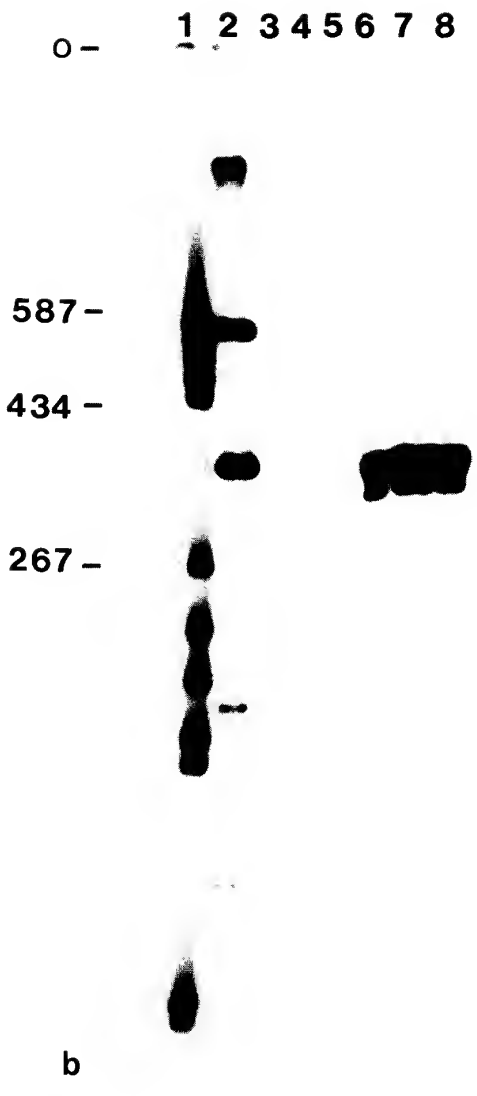


Fig. 2b.

640 and 778, and 683 and 743, respectively.

Hood and his colleagues have recently determined the sequence of 35 amino-terminal amino acids of human lymphoblastoid interferon (Zoon et al., 1980; M. Hunkapiller and L. Hood, personal communication). In Fig. 5 the sequence of human lymphoblastoid IF (B) is aligned with the amino acid sequence determined by the major reading frame of the Hif-2h nucleotide sequence (A) such that the amino-terminal acid of the former coincides with the amino acid coded for by the 24th codon of the latter. Extensive coincidence is found: 26 of 35 positions have identical amino acids. This confirms the assignment of the reading frame.

DISCUSSION

Cloned cDNA generated from poly(A) RNA by commonly used procedures (Efstratiadis et al., 1977) lacks 5′-terminal nucleotides and may even contain artifactual sequences (Richards et al., 1979). It is therefore not certain whether the first AUG of the cloned human LeIF cDNA Hif-2h, which is located 57 nucleotides downstream from the 5′-terminus of the heteropolymeric sequence, in fact corresponds to the first AUG on the mRNA. Bearing these reservations in mind, we shall assume, until further experimental evidence becomes available, that this is the case.

In eukaryotic mRNAs, the first AUG triplet from the 5′-terminus is usually the initiation site for protein synthesis (Kozak, 1978). The codon in the cloned human LeIF cDNA corresponding to the first amino acid of lymphoblastoid interferon is 22 codons downstream from the first AUG (and 14 codons downstream from the second one) indicating that the sequence coding for interferon may be preceded by a sequence determining a signal peptide of 23 (or less likely 15) amino acids (Fig. 4). The longer of the presumptive signal sequences contains an uninterrupted series of 11 hydrophobic amino acids (and the shorter one, one of 6). This accumulation of hydrophobic residues is characteristic of signal sequences (see Davis and Tai, 1980). The presumptive cleavage site between signal and interferon sequence lies between a Gly and a Cys residue. It is striking that in the case of *E. coli* prelipoprotein, cleavage occurs between the same two amino acids (Inouye et al., 1977). It will be interesting to determine whether the postulated preinterferon exists, and if so, whether it is correctly processed in *E. coli*, especially in view of our finding (S. Nagata, unpublished results) that about 50% of the interferon activity produced in *E. coli* can be released by osmotic shock and is therefore located in the periplasmic space (Anraku, 1968).

The sequence corresponding to (mature) LeIF polypeptide comprises 498 nucleotides, which code

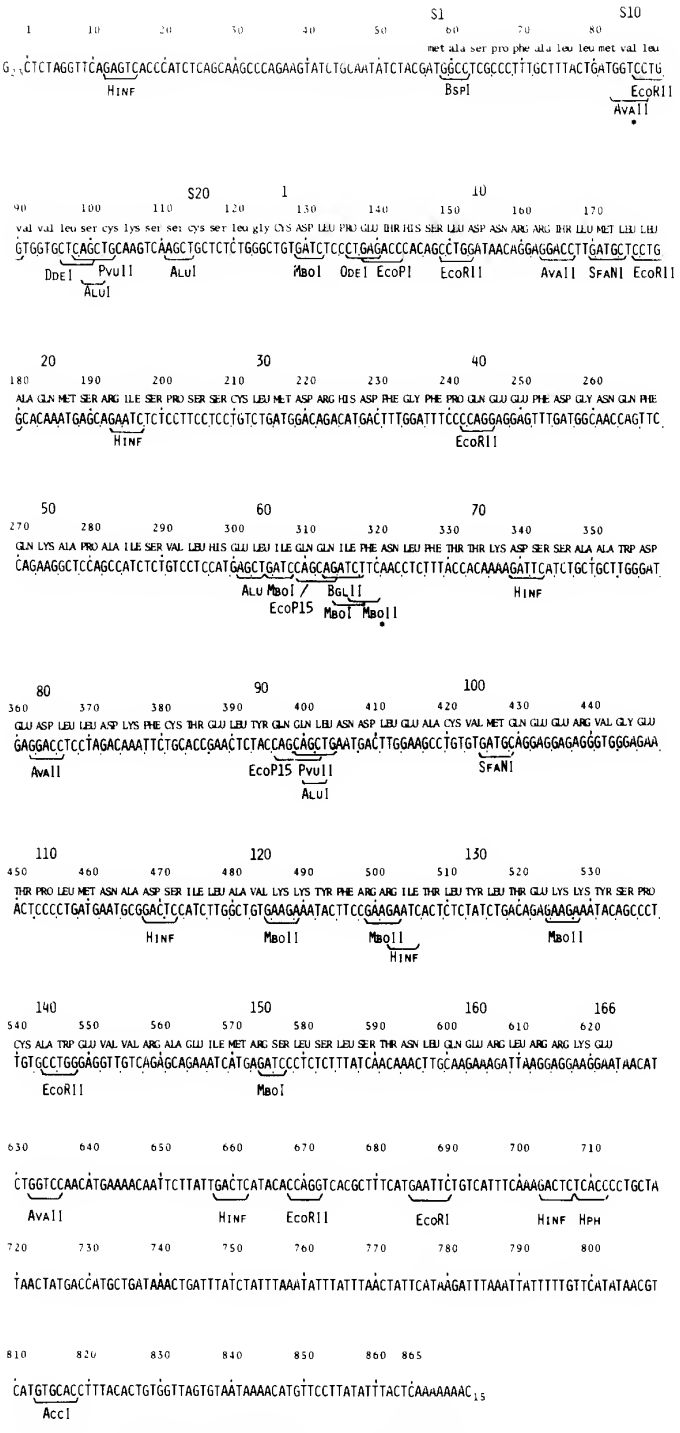


Fig. 3. The nucleotide sequence of Hif-2h IF cDNA. The nucleotide sequence was determined as indicated in METHODS. The amino acid sequence was deduced from the nucleotide sequence; lower case letters indicate the putative signal polypeptide The *Mbo*I sites as well as the restriction targets marked with * were resistant to cleavage.

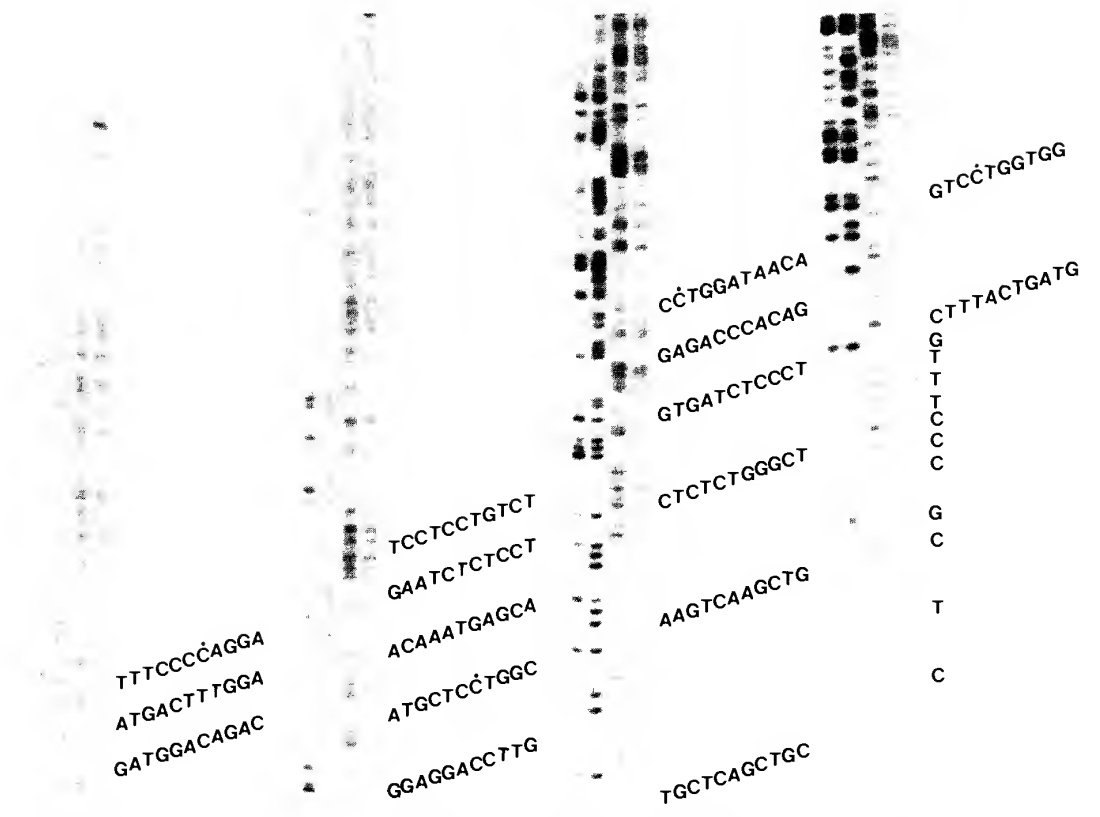


Fig. 4. Autoradiogram of the sequence gel showing the presumed signal polypeptide coding region and the beginning of the interferon polypeptide. Fragment b (Fig. 1) was degraded as described in METHODS and analyzed on a 12% gel. The runs were for 2, 8, 18 and 26 h at 900 V. The four lanes for each run show, from left to right, degradations specific for G, A + G, C + T and C. Because of methylation at *Eco*RII sites, the Cs marked with a dot above them are present as gaps in this sequence. Their existence was confirmed by analyses of the other strand.

for 166 amino acids. Assuming that there is no carboxyterminal processing, the molecular weight of the interferon polypeptide, as calculated from Table I, is 19388. The base composition of the coding sequence is 50% G + C; the codon usage within the interferon coding sequence (Table II) is in reasonable agreement with that compiled for mammalian mRNAs in general (Grantham et al., 1980); the deviations observed may be ascribed to the small numbers involved.

The non-coding 3′-region consists of 242 nucleotides; this length is intermediate between that of chicken ovalbumin mRNA (637 residues) (McReynolds et al., 1978) and rat insulin mRNA (53) (Ullrich et al., 1977). The high A + T content (69%) is similar to that found for the corresponding segment of mouse β-globin minor mRNA (63%) (Konkel et al., 1979); the A + T content of eukaryotic non-coding 3′-regions range from 94% in mRNA yeast mitochondrial ATPase (Hensgens et al., 1979) to 42% in bovine ACTH-βLPH mRNA (Nakanishi et al., 1979). No striking homologies to non-coding 3′-regions of other mRNAs were noted, except for the AATAAA-(AC) sequence 18–27 nucleotides upstream from the poly(A) sequence, found previously (Proudfoot and Brownlee, 1976) in almost all eukaryotic mRNAs examined, at about the same relative position.

The comparison of the first 35 amino acids of lymphoblastoid interferon (Zoon et al., 1980; M. Hunkapiller and L. Hood, personal communication) and the sequence deduced from Hif-2h (Fig. 5) shows 9 differences. In all cases, the codons for the differing amino acids could be related by one-base changes. The amino acid compositions (Table I) determined directly for lymphoblastoid interferon on

TABLE I

Amino acid composition of leukocyte and lymphoblastoid interfcron (IF) [a]

	Number of amino acid residues		
	Leukocyte IF, deduced from nucleotide sequence of Hif-2h cDNA [b]	Lymphoblastoid IF [c]	LeIF
Asn	6	15	15
Asp	11		
Thr	9	8	7.5
Ser	13	11	8
Gln	10	27	24
Glu	15		
Pro	*6*	*11*	6
Gly	*3*	*11*	5.5
Ala	10	11	8
Cys	*5*	*2*	3
Val	6	8	8
Met	*6*	*3* *	4
Ile	7	7	9
Leu	22	18	22
Tyr	4	4	5
Phe	8	7	9
His	3	4	3
Lys	8	10	12
Arg	12	10	7
Trp	2	1	1

[a] Values believed to differ significantly are italicized.
[b] From Fig. 3.
[c] From Zoon et al. (1980), except for the value marked with *, which was from L. Hood (personal communication).
[d] From Rubinstein et al. (1979).

A) CYS ASP LEU PRO GLU THR HIS SER LEU ASP ASN ARG ARG THR LEU
B) SER - - - GLN - - - - GLY - - - ALA -

A) MET LEU LEU ALA GLN MET SER ARG ILE SER PRO SER SER CYS LEU
B) ILE - - - - - GLY - - - LEU PHE - - -

A) MET ASP ARG HIS ASP
B) LYS - - - -

Fig. 5. Comparison of the amino terminal amino acid sequence of lymphoblastoid interferon (determined experimentally) and leukocyte interferon (deduced from the Hif-2h cDNA nucleotide sequence). The leukocyte interferon sequence (A) is from Fig. 3; the lymphoblastoid interferon sequence (B) is from Zoon et al. (1980), and M. Hunkapiller and L. Hood (personal communication). Dashes indicate identical amino acids.

TABLE II

Codon usage in the interferon coding sequence [a]

Codon		all MAM	IG	MAM-IG	LeIF
Arg	CGA	2	7	0	6
	CGC	9	4	12	0
	CGG	6	3	8	0
	CGU	7	1	9	0
	AGA	6	12	3	36
	AGG	11	9	11	30
Leu	CUA	8	11	7	6
	CUC	24	26	24	48
	CUG	51	16	68	42
	CUU	7	7	7	0
	UUA	4	11	1	12
	UUG	7	10	6	24
Ser	UCA	11	24	5	12
	UCC	20	17	22	24
	UCG	4	1	5	0
	UCU	17	29	11	24
	AGC	21	27	18	18
	AGU	17	33	9	0
Thr	ACA	13	26	7	18
	ACC	25	30	23	24
	ACG	7	4	9	0
	ACU	19	36	11	12
Pro	CCA	12	20	7	6
	CCC	18	13	21	12
	CCG	8	5	9	0
	CCU	13	13	12	18
Ala	GCA	11	20	6	12
	GCC	34	22	40	18
	GCG	5	1	8	6
	GCU	23	21	24	24
Gly	GGA	10	20	5	12
	GGC	28	16	34	6
	GGG	11	9	11	0
	GGU	18	25	15	0
Val	GUA	3	5	1	0
	GUC	16	23	13	12
	GUG	35	17	44	18
	GUU	7	7	7	6
Lys	AAA	15	17	14	24
	AAG	45	24	56	24
Asn	AAC	29	29	29	24
	AAU	10	13	9	12
Gln	CAA	10	11	9	12
	CAG	32	30	32	48
His	CAC	22	8	29	6
	CAU	10	11	10	12

TABLE II (continued)

Codon		all MAM	IG	MAM–IG	*Le* IF
Glu	GAA	22	21	23	36
	GAG	36	23	43	54
Asp	GAC	27	22	30	36
	GAU	18	23	16	30
Tyr	UAC	20	20	20	18
	UAU	15	17	14	6
Cys	UGC	12	6	16	6
	UGU	11	18	8	24
Phe	UUC	33	30	35	24
	UUU	16	13	18	24
Ile	AUA	4	7	3	0
	AUC	20	21	19	42
	AUU	11	18	8	0
Met	AUG	16	14	16	36
Trp	UGG	16	25	12	12

[a] The values are expressed per 1000 amino acid residues. The data for "all MAM" (compiled from 18 mammalian sequences), "IG" (6 immunoglobulin sequences) and "MAM–IG" (12 non-immunoglobulin mammalian sequences) are from Grantham et al. (1980), those for human leukocyte interferon are from Fig. 3.

the one hand, and deduced from the Hif-2h sequence, on the other, show striking differences in regard to their content of Gly, Pro, Cys and Met. These differences are too large to be explained by polymorphism; most likely we are dealing with the products of two non-allelic genes, since the degree of divergence of the two proteins (26% mismatch) is similar to that between, for example, human and sheep β-globin (23% mismatch). We have recently surveyed our human leukocyte cDNA clone bank and identified a hybrid plasmid which also directs synthesis of interferon activity in *E. coli* and has a different restriction pattern than Hif-2h (M. Streuli and M. Schwarzstein, unpublished results). This clone represents a second leukocyte interferon gene (LeIF II), differing from the one (LeIF I) which corresponds to Hif-2h. The amino acid composition of an IF preparation from human leukocytes (Rubinstein et al., 1979) agrees somewhat better than that of lymphoblastoid IF with the amino acid composition deduced for LeIF I (Table I). This clone represents a second leukocyte interferon gene.

Taniguchi and his colleagues prepared cDNA from induced fibroblast poly(A) RNA and selected presumptive interferon cDNA clones by hybridization techniques (Taniguchi et al., 1979). The nucleotide sequence of one such clone was determined and could be correlated (Taniguchi et al., 1980) with the sequence of the 13 amino terminal amino acids of fibroblast interferon (Knight et al., 1980). The striking structural homologies between the leukocyte and fibroblast interferon cDNA sequences will be analyzed elsewhere.

ACKNOWLEDGEMENTS

We thank A. van Ooyen for helpful suggestions concerning the sequencing, W. Boll and A. Kiss for restriction enzymes, J. Ecsödi for plasmid DNA and restriction analyses and M. Hunkapiller, L. Hood and T. Taniguchi for unpublished information. S.P. is a Scholar of the Leukemia Society of America. The work was supported by Biogen S.A. and the Schweizerische Nationalfonds.

REFERENCES

Anraku, Y.: Transport of sugars and amino acids in bacteria. J. Biol. Chem. 243 (1968) 3116–3122.

Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heyneker, H.L. and Boyer, H.W.: Construction and characterization of new cloning vehicles, II. A multipurpose cloning system. Gene 2, (1977) 95–113.

Davis, B.D. and Tai, P.-C.: The mechanism of protein secretion across membranes. Nature 283 (1980) 433–438.

Efstratiadis, A., Kafatos, F.C. and Maniatis, T.: The primary structure of rabbit β-globin mRNA as determined from cloned DNA. Cell 10 (1977) 571–588.

Grantham, R., Gautier, C., Gouy, M., Mercier, R. and Pave, A.: Codon catalog usage and the genome hypothesis. Nucl. Acids Res. 8 (1980) r49–r62.

Hensgens, L.A.M., Grivell, L.A. Borst, P. and Bos, J.L.: Nucleotide sequence of the mitochondrial structural gene for subunit 9 of yeast ATPase complex. Proc. Natl. Acad. Sci. USA 76 (1979) 1663–1667.

Inouye, S., Wang, S., Sekizawa, J., Halegoua, S. and Inouye, M.: Amino acid sequence for the peptide extension on the prolipoprotein of the *Escherichia coli* outer membrane. Proc. Natl. Acad. Sci. USA 74 (1977) 1004–1008.

Konkel, D.A., Maizel, J.V., Jr. and Leder, P.: The evolution and sequence comparison of two recently diverged mouse chromosomal β-globin genes. Cell 18 (1979) 865–873.

Kozak, M.: How do eukaryotic ribosomes select initiation regions in messenger RNA? Cell 15 (1978) 1109–1123.

Loening, U.E.: The fractionation of high-molecular weight ribonucleic acid by polyacrylamide gel electrophoresis. Biochem. J. 102 (1967) 251–257.

Maxam, A.M. and Gilbert, W.: A new method for sequencing DNA. Proc. Natl. Acad. Sci. USA 74 (1977) 560–564.

McReynolds, L., O'Malley, B.W., Nisbet, A.D., Fothergill, J.E., Givol, D., Fields, S., Robertson, M. and Brownlee, G.G.: Sequence of chicken ovalbumin mRNA. Nature 273 (1978) 723–728.

Müller, W., Weber, H., Meyer, F. and Weissmann, C.: Site-directed mutagenesis in DNA: Generation of point mutations in cloned β globin complementary DNA at the positions corresponding to amino acids 121 to 123. J. Mol. Biol. 124 (1978) 343–358.

Nagata, S., Taira, H., Hall, A., Johnsrud, L., Streuli, M., Ecsödi, J., Boll, W., Cantell, K. and Weissmann, C.: Synthesis in *E. coli* of a polypeptide with human leukocyte interferon activity. Nature 284 (1980) 316–320.

Nakanishi, S., Inoue, A., Kita, T., Nakamura, M., Chang, A.C.Y., Cohen, S.N. and Numa, S.: Nucleotide sequence of cloned cDNA for bovine corticotropin-β-lipotropin precursor. Nature 278 (1979) 423–427.

Proudfoot, N.J. and Brownlee, G.G.: 3′ non-coding region sequences in eukaryotic messenger RNA. Nature 263 (1976) 211–214.

Richards, R.I., Shine, J., Ullrich, A., Wells, J.R.E. and Goodman, H.M.: Molecular cloning and sequence analysis of adult chicken β globin cDNA. Nucl. Acids Res. 7 (1979) 1137–1146.

Rubinstein, M., Rubinstein, S., Familletti, P.C., Miller, R.S., Waldman, A.A. and Pestka, S.: Human leukocyte interferon: Production, purification to homogeneity, and initial characterization. Proc. Natl. Acad. Sci. USA 76 (1979) 640–644.

Smith, H.O. and Birnstiel, M.L.: A simple method for DNA restriction site mapping. Nucl. Acids Res. 3 (1976) 2387–2398.

Sutcliffe, J.G.: pBR322 restriction map derived from the DNA sequence, accurate DNA size markers up to 4361 nucleotide pairs long. Nucl. Acids Res. 5 (1978) 2721–2728.

Taniguchi, T., Sakai, M., Fujii-Kuriyama, Y., Muramatsu, M., Kobayashi, S. and Sudo, T.: Construction and identification of a bacterial plasmid containing the human fibroblast interferon gene sequence. Proc. Jap. Acad. Ser. B 55 (1979) 464–469.

Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. and Muramatsu, M.: The nucleotide sequence of human fibroblast interferon cDNA. Gene 10 (1980) 11–15.

Ullrich, A., Shine, J., Chirgwin, J., Pictet, R., Tischer, E., Rutter, W.J. and Goodman, H.M.: Rat insulin gene: construction of plasmids containing the coding sequences. Science 196 (1977) 1313–1319.

Weaver, R.F. and Weissmann, C.: Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15 S β-globin mRNA precursor and mature 10 S β-globin mRNA have identical map coordinates. Nucl. Acids Res. 7 (1979) 1175–1193.

Wiegand, R.C., Godson, G.N. and Radding, C.M.: Specificity of the S_1 nuclease from *Aspergillus oryzae*. J. Biol. Chem. 250 (1975) 8848–8855.

Wilkie, N.M., Clements, J.B., Boll, W., Mantei, N., Londsdale, D. and Weissmann, C.: Hybrid plasmids containing an active thymidine kinase gene of *Herpes simplex* virus 1. Nucl. Acids Res. 7 (1979) 859–877.

Zoon, K.C., Smith, M.E., Bridgen, P.J., Anfinsen, C.B., Hunkapiller, M.W. and Hood, L.E.: Amino terminal sequence of the major component of human lymphoblastoid interferon. Science 207 (1980) 527–528.

Communicated by W. Szybalski.

TAB G

40. Sutcliffe, J. G. *Nucleic Acids Res.* 5, 2721-2728 (1978).
41. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 74, 560-564 (1977).
42. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. O. & Goodman, H. M. *Nature* 270, 486-494 (1977).
43. Proudfoot, N. J. & Brownlee, G. G. *Nature* 263, 211-214 (1976).
44. Fiers, W. *et al. Nature* 273, 113-120 (1978).
45. Soeda, E., Arrand, J. R., Smolar, N., Walsh, J. E. & Griffin, B. E. *Nature* 283, 445-453 (1980).
46. Min Jou, W. *et al. Cell* 19, 683-696 (1980).
47. Shine, J., Seeburg, P. H., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* 270, 494-499 (1977).
48. Blobel, G. *et al. Soc. exp. Biol. Symp.* 33, 9-36 (1979).
49. Habener, J. F. *et al. Proc. natn. Acad. Sci. U.S.A.* 75, 2616-2620 (1978).
50. Oavey, M. W., Sulkowski, E. & Carter, W. A. *J. biol. Chem.* 251, 7620-7625 (1976).
51. Neuberger, A., Gottschalk, A., Marshall, R. D. & Spiro, R. G. in *The Glycoproteins: Their Composition, Structure and Function* (ed. Gottschalk, A.) 450-490 (Elsevier, Amsterdam, 1972).
52. Tomita, M. & Marchesi, V. *Proc. natn. Acad. Sci. U.S.A.* 72, 2964-2968 (1975).
53. Taniguchi, T. *et al. Gene* 10, 11-15 (1980).
54. Taniguchi, T. *et al. Nature*, 285, 547-549 (1980).
55. Mantei, N. *et al. Gene* 10, 1-10 (1980).
56. Roberts, R. J. *Gene* 8, 329-343 (1980).

Human leukocyte and fibroblast interferons are structurally related

Tadatsugu Taniguchi*, Ned Mantei†, Marco Schwarzstein†, Shigekazu Nagata†, Masami Muramatsu* & Charles Weissmann†

*Department of Biochemistry, Cancer Institute, Japanese Foundation for Cancer Research, Tokyo 170, Japan

†Institut für Molekularbiologie I, Universität Zürich, 8093 Zürich, Switzerland

The coding sequences of the cDNAs of cloned human leukocyte interferon I and human fibroblast interferon show homologies of 45% at the nucleotide and 29% at the amino acid level. We conclude that the two genes were derived from a common ancestor.

THE acid-stable human interferons are subdivided into two major groups, fibroblast interferons (F-IF) and leukocyte interferons (Le-IF); these are the major components of the interferons produced by induced fibroblasts and leukocytes, respectively. Some cells, such as the lymphoblastoid Namalva cell line, produce a mixture of 90% Le-IF and 10% F-IF[1,2]. The two interferon types have several features in common: both are glycoproteins with molecular weights ranging from 16,000 to 26,000 (refs 3–9), the induction and shut-off of their synthesis seem to be under similar control[6], and at least some of the responses elicited in target cells are similar, such as induction of an antiviral state, which is accompanied by increased synthesis of several proteins[10–13]. Nonetheless, the two kinds of interferon differ in many respects. Antibodies directed against Le-IF do not neutralize F-IF and vice versa[14], the target cell specificities of the two interferons differ[15], and the sequences of the 13 amino-terminal amino acids of F-IF and Le-IF (from lymphoblastoid cells) show no homology[16,17]. Although Le-IF and F-IF are encoded by different mRNA species[18], it is not known whether these mRNAs are transcribed from distinct genes or whether they arise from the same gene through a common precursor which is processed or spliced in different modes.

We have recently cloned and sequenced one species each of Le-IF (Le-IF I)[19,20] and F-IF cDNA[21,22]. A second species of Le-IF (Le-IF II) cDNA has recently been identified (M. Streuli, S.N. and C.W., unpublished results).

Comparison of the amino acid sequences of Le-IF and F-IF

In Fig. 1 the nucleotide sequences of Le-IF I and F-IF cDNA were aligned so that the AUGs closest to their 5′ termini coincided. From the amino-terminal sequence published for F-IF[16] and lymphoblastoid Le-IF[17], one can determine that for F-IF and Le-IF, respectively, the 21st and 23rd codons following the initiation triplet represent the first amino acid of the interferon polypeptide. Presumably, the stretch in between encodes a signal peptide. As the respective putative signal peptides of Le-IF and F-IF comprise 23 and 21 amino acids, the IF polypeptides, as aligned in Fig. 1, are shifted by two residues relative to their termini. In this alignment, 48 of 166 positions (29%) of the interferon polypeptides have identical amino acids. By introducing appropriate gaps, better homology could be achieved, particularly in the region of the signal sequence; in the present comparison this has not been done.

To plot the degree of homology between the F-IF and Le-IF as a function of the map distance, the sequence was subdivided into segments of 8 amino acids (or 24 nucleotides), each overlapping by 4 amino acids (or 12 nucleotides) with the neighbouring segments, and the per cent coincidence of amino acids (and nucleotides) for each segment was determined (see ref. 23). As seen in Fig. 2, amino acid sequences show three domains of homology. The first one, with the least degree of homology, corresponds to the putative signal sequence, which is rich in hydrophobic residues and has 4 identical amino acid positions out of 21; the second domain, between amino acids 28 and 80 (counted on the Le-IF sequence), has 21 identical residues out of 51 (41% homology), and the third, between positions 115 and 151 (Le-IF sequence), has 19 out of 35 identical residues (54%). The longest stretches of contiguous conserved amino acids are Gln-Phe-Gln-Lys (positions 47–50 of Le-IF and 49–52 of F-IF) and Cys-Ala-Trp (positions 139–141 and 141–143, respectively); the latter sequence is notable because it comprises Cys and Trp, which are preferentially conserved in related proteins[24]. Table 1 shows that conservation was highest between the interferon polypeptides (not considering the signal sequences) for Trp, Phe, Arg, Cys and Tyr residues, in agreement with the general experience that the amino acids most likely to be conserved between related proteins are Trp > Cys > Tyr > Arg > Phe, His (ref. 24). Even where amino acids are conserved, the codons show one or more nucleotide changes in half the instances. The codons of three out of seven conserved Leu residues are unrelated, as are two of four codons pertaining to conserved Ser residues. This suggests that there is a strong selective pressure favouring the conservation of several amino acids. It is quite likely that at least some of the conserved amino acids are essential for a function common to Le-IF and F-IF, perhaps the induction of the virus-resistant state in the target cell. These findings may provide guidelines for the tailoring of modified[25], possibly shorter polypeptides possessing certain activities of interferon.

Comparison of the nucleotide sequences of Le-IF and F-IF

The nucleic acid sequences show an average homology of 43% in the domain of the signal sequence and of 45% in the interferon polypeptide sequence. On a random basis, about

0028-0836/80/250547-03$01.00

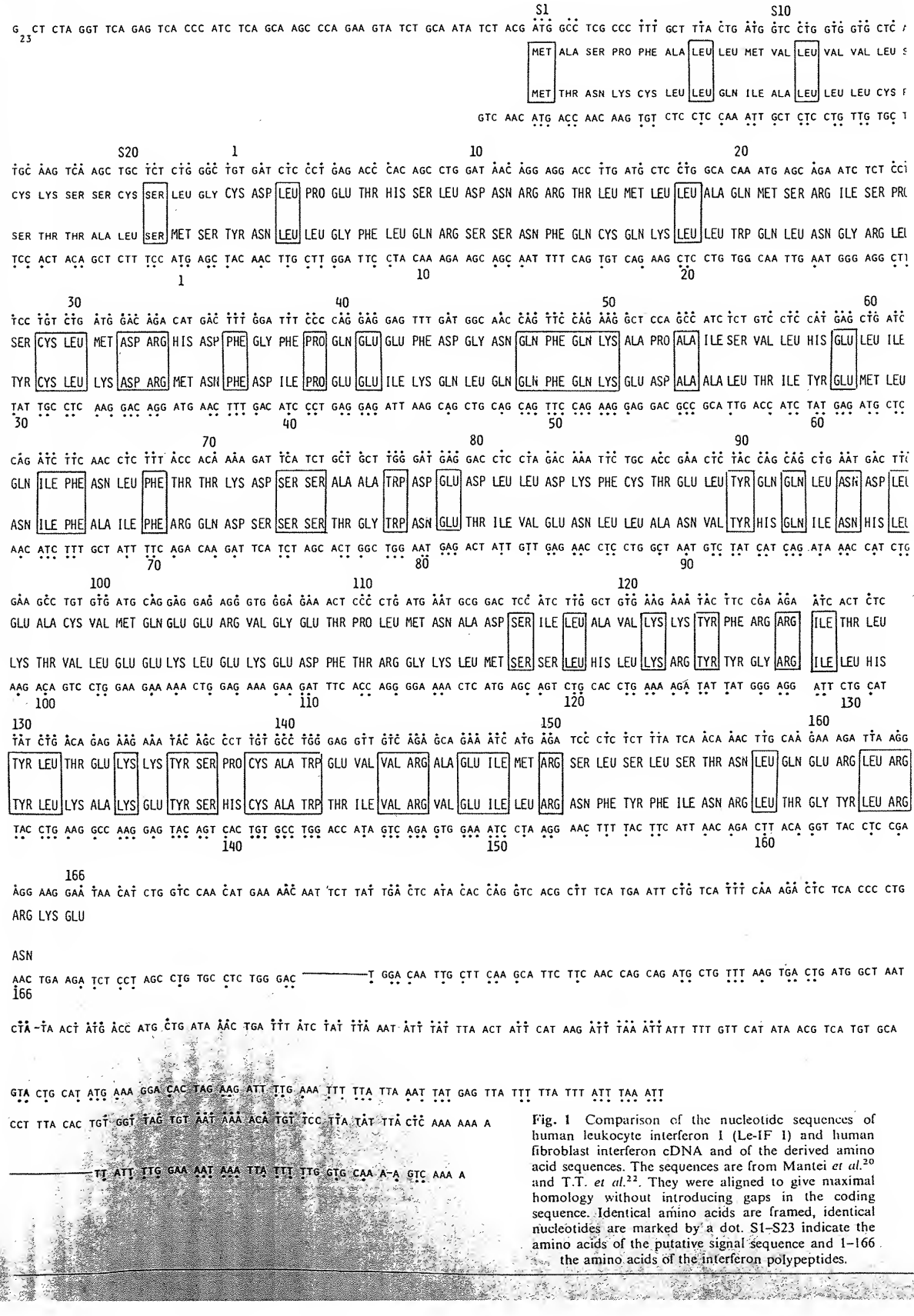


Fig. 1 Comparison of the nucleotide sequences of human leukocyte interferon 1 (Le-IF 1) and human fibroblast interferon cDNA and of the derived amino acid sequences. The sequences are from Mantei *et al.*[20] and T.T. *et al.*[22]. They were aligned to give maximal homology without introducing gaps in the coding sequence. Identical amino acids are framed, identical nucleotides are marked by a dot. S1–S23 indicate the amino acids of the putative signal sequence and 1–166 the amino acids of the interferon polypeptides.

25% of the nucleotide positions should coincide. Within the interferon coding sequence, the nucleotide homologies are more evenly distributed than the amino acid homologies. However, one may distinguish, albeit to a less pronounced degree, the same three blocks of similarity noted for the amino acids. The longest region without mismatches extends for 13 nucleotides (compare 47th to 51st codon of Le-IF with 49th to 53rd codon of F-IF). There are, in addition, sequences of 17, 18 and 20 nucleotides with 3, 3 and 4 mismatches, respectively. The heteropolymeric 3′-terminal noncoding region of Le-IF cDNA has 242 nucleotides, and is longer by 39 residues than its counterpart in F-IF cDNA. In aligning the two sequences, four gaps were introduced to maximize homology, as described by van Ooyen *et al.*[23]. In this way, several segments were matched with 29–41% homology. The introduction of gaps in the alignment may be justified in view of the arguments presented previously, that intervening sequences and noncoding regions of reduplicated genes diverge as a consequence of block insertions and/or deletions in the course of evolution[23,26]. It is unlikely that the extent of homology between Le-IF and F-IF cDNA would allow significant cross-hybridization between the two species.

A common ancestral gene for Le-IF and F-IF

On the basis of our findings, there is no doubt that Le-IF and F-IF genes are derived from a common ancestral sequence. When did the separation of these genes occur? Human α- and β-globin show 57% amino acid mismatches, and human β-globin and myoglobin, as well as α-globin and myoglobin, 76% mismatches. If the rate of divergence of interferons and globins is comparable (however, this is quite uncertain, see ref. 24, p. 50, for proteins showing both higher and lower rates), the separation of interferon genes occurred after that of myoglobin and haemoglobins but before that of α- and β-globins, that is between 500 and 1,000 Myr ago[24], which is about the time vertebrates arose[27]. This would mean that both types of interferon gene should occur in all vertebrates, unless one and/or the other was lost by deletion. Indeed, as shown by the sequencing of 13–24 amino-terminal residues[16,17,28], mouse interferons A and B show significant homology with human fibroblast interferon, and mouse interferon C with human lymphoblastoid interferon, whereas the mouse species A and B on the one hand, and the species C on the other show no significant homology within the short segment sequenced. Thus, at least in the mouse, representatives of both interferon families exist. It will be of interest to determine the evolutionary relationship of these to the third type of interferon, immune or γ-interferon.

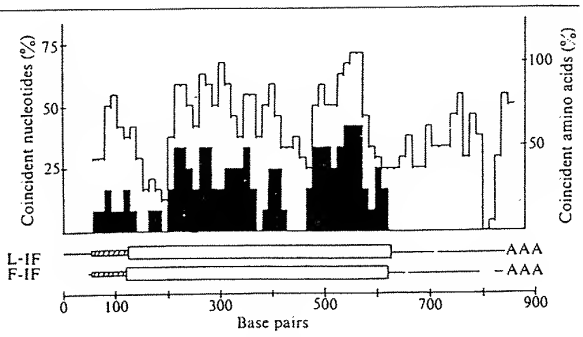


Fig. 2 Similarity of the nucleotide and amino acid sequences of human leukocyte interferon I and fibroblast interferon. The sequences shown in Fig. 1 were subdivided into segments of 8 amino acids or 24 nucleotides, each overlapping by 4 and 12 residues, respectively, with the neighbouring segments. The percentage of concident residues was plotted as a function of map position. Open vertical blocks, nucelotides; filled vertical blocks, amino acids. L-IF, leukocyte interferon cDNA; F-IF, fibroblast interferon cDNA; lines, noncoding sequences; hatched bars, putative signal peptide; open bars, interferon polypeptide.

After submission of this article, we learnt that Derynck *et al.* had cloned and sequenced fibroblast interferon (see accompanying article[29]), confirming the deduced amino acid sequence of T.T *et al.*[22].

T.T. and M.M. thank Dr H. Sugano for his continuous support and interest. The work at the University of Zürich was supported by Biogen SA and the Schweizerische Nationalfonds.

Table 1 Conservation of amino acids in leukocyte and fibroblast interferon

	F-IF	Le-IF	Conserved amino acids	No. of changes in codon of conserved amino acids 0	1	2	3
Leu	25	22	8	1	4	3	
Cys	3	5	2	1	1		
Asn	12	6	1	1			
Arg	11	12	5	1	3	1	
Phe	9	8	4	2	2		
Pro	1	6	1		1		
Gln	11	10	3	3			
Lys	11	8	3	2	1		
Ala	6	10	2	2			
Glu	13	15	4	4			
Ile	11	7	3	2	1		
Ser	9	13	4		2	1	1
Trp	3	2	2	2			
Tyr	10	4	4	1	3		
Val	5	6	1	1			
Asp	5	11	1	1			
Thr	6	9	0				
Gly	6	3	0				
Met	4	6	0				
His	5	3	0				
Total	166	166	48	24	18	5	1

The data are from T.T. *et al.*[22] and Mantei *et al.*[20].

Received 24 March; accepted 29 April 1980.

1. Havell, E. A., Yip, Y. K. & Vilček, J. *J. gen. Virol.* 38, 51–59 (1977).
2. Paucker, K. *Tex. Rep. Biol. Med.* 35, 23–28 (1977).
3. Bridgen, P. J. *et al. J. biol. Chem.* 252, 6585–6587 (1977).
4. Knight, E. Jr *Proc. natn. Acad. Sci. U.S.A.* 73, 520–523 (1976).
5. Tan, Y. H. *et al. J. biol. Chem.* 254, 8067–8073 (1979).
6. Hayes, T. G., Yip, Y. K. & Vilček, J. *Virology* 98, 351–363 (1979).
7. Rubinstein, M. *et al. Proc. natn. Acad. Sci. U.S.A.* 76, 640–644 (1979).
8. Stewart, W. E. II & Wiranowska-Stewart, M. *Virology* 97, 473–476 (1979).
9. Havell, E. A., Yamazaki, S. & Vilček, J. *J. biol. Chem.* 252, 4425–4427 (1977).
10. Lebleu, B., Sen, G. C., Shaila, S., Cabrer, B. & Lengyel, P. *Proc. natn. Acad. Sci. U.S.A.* 73, 3107–3111 (1976).
11. Hovanessian, A. G., Brown, R. E. & Kerr, I. M. *Nature* 268, 537–540 (1977).
12. Hovanessian, A. G. & Kerr, I. M. *Eur. J. Biochem.* 93, 515–526 (1979).
13. Schmidt, A. *et al. Proc. natn. Acad. Sci. U.S.A.* 76, 4788–4792 (1979).
14. Havell, E. A. *et al. Proc. natn. Acad. Sci. U.S.A.* 72, 2185–2187 (1975).
15. Stewart, W. E. II in *The Interferon System*, 134–145 (Springer, New York, 1979).
16. Knight, E. Jr., Hunkapiller, M. W., Korant, B. O., Hardy, R. W. F. & Hood, L. E. *Science* 207, 525–526 (1980).
17. Zoon, K. C. *et al. Science* 207, 527–528 (1980).
18. Cavalieri, R. L., Havell, E. A., Vilček, J. & Pestka, S. *Proc. natn. Acad. Sci. U.S.A.* 74, 3287–3291 (1977).
19. Nagata, S. *et al. Nature* 284, 316–320 (1980).
20. Mantei, N. *et al. Gene* 10, 1–10 (1980).
21. Taniguchi, T. *et al. Proc. Jap. Acad.* B55, 464–469 (1979).
22. Taniguchi, T., Ohno, S., Fujii-Kuriyama, Y. & Muramatsu, M. *Gene* 10, 11–15 (1980).
23. van Ooyen, A., van den Berg, J., Mantei, N. & Weissmann, C. *Science* 206, 337–344 (1979).
24. Dayhoff, M. O. in *Atlas of Protein Sequence and Structure* Vol. V (Natn. Biomed. Res. Fdn, Washington, 1972).
25. Müller, W., Weber, H., Meyer, F. & Weissmann, C. *J. molec. Biol.* 124, 343–358 (1978).
26. Konkel, D. A., Maizel, J. V. Jr & Leder, P. *Cell* 18, 865–873 (1979).
27. Kurtén, B. in *Proc. XVI int. Congr. Zool.* Vol. 6 (ed. Moore, J. A.) 327–354 (Natural History Press, New York, 1965).
28. Taira, H. *et al. Science* 207, 528–530 (1980).
29. Derynck, R. *et al. Nature* 285, 542–547 (1980).

Filed on behalf of Party **Goeddel**

By: Thomas E. Friebel
Michael J. Ryan
Jones Day
222 East 41st Street
New York, New York 10017-6702
Telephone: (212) 326-3939
Fax: (212) 755-7306
tefriebel@jonesday.com
mjryan@jonesday.com

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE BOARD OF PATENT APPEALS
AND INTERFERENCES

HARUO **SUGANO**
MASAMI MURAMATSU, and TADATSUGU TANIGUCHI

Junior Party
(Patent 5,514,567 and 5,326,859; Application 08/463,757)

v.

DAVID V. **GOEDDEL**
and ROBERTO CREA

Senior Party
(Application 07/374,311; Patent 5,460,811)

Patent Interference Nos. **105,334** and **105,337** (SGL)
(Technology Center 1600)

SECOND DECLARATION OF RIK DERYNCK, Ph.D.

Sugano Exhibit 1012
Fiers v. Sugano
Interference 105,661

Goeddel EXHIBIT 1108
Interference 105,334
Interference 105,337

SECOND DECLARATION OF RIK DERYNCK, Ph.D.

I, RIK DERYNCK, declare and state as follows:

1. I reside at 27 Hilltop Road, San Mateo, California 94402.

2. I received a Ph.D. (summa cum laude) in Molecular Biology from the University of Ghent, Belgium, in 1981, where I was a Predoctoral Fellow/Research Associate with Dr. Walter Fiers.

3. Since April, 1991 I have been Professor, first in the Department of Growth and Development and subsequently in the Department of Cell and Tissue Biology, University of California at San Francisco, and I am now Co-Director of the UCSF Institute for Regeneration Medicine, and Director of its Program in Craniofacial and Mesenchymal Biology.

4. I have conducted research in molecular biology and recombinant DNA technology prior to and since receiving my Ph.D.

5. Prior to joining the University of California at San Francisco, I was Senior Scientist (March 1985 to December 1990) in the Departments of Molecular and Developmental Biology and Scientist (April 1981 to March 1985) in the Department of Molecular Biology at Genentech Inc., South San Francisco, California.

6. I have authored or co-authored 209 publications in the fields of molecular biology and developmental biology and received a number of honors and awards, which are set forth in my curriculum vitae (Exhibit 1028).

7. During my time with the laboratory of Dr. Fiers in Ghent, Belgium, I was the lead investigator in research culminating in the cloning of cDNA encoding the precursor form of human fibroblast interferon. I was also the lead investigator in the research that resulted in the successful expression in *E. coli* of fusion proteins having the amino acids of the precursor form

of human fibroblast interferon with additional amino acids attached to its amino terminus. In addition, I played a major role in the expression of mature human fibroblast interferon having a total of 166 amino acids and unaccompanied by its signal peptide.[1] This research was reported in a number of publications, particularly the following:

(a) Derynck *et al.*, "Isolation and structure of a human fibroblast interferon gene," Nature 285: 542-547 (1980) (Exhibit 1010)

(b) Derynck *et al.*, "Expression of human fibroblast interferon gene in *Escherichia coli*," Nature 287: 193-197 (1980) (Exhibit 1011)

(c) Remaut *et al.*, "Inducible high level synthesis of mature human fibroblast interferon in *Escherichia coli*," Nucl. Acids Res. 11: 4677-4688 (1983) (Exhibit 1029).

8. I am familiar with the state of the art as of March 19, 1980 with respect to the cloning and expression of eukaryotic genes in bacteria, particularly *E. coli*. In this context, I have reviewed and considered the following publications:

(a) Ratzkin and Carbon, "Functional expression of cloned yeast DNA in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 74: 487-491 (1977) (Exhibit 2078)

(b) Vapnek *et al.*, "Expression in *Escherichia coli* K-12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*," Proc. Nat'l Acad. Sci. USA 74: 3508-3512 (1977) (Exhibit 2079)

[1] The terms signal peptide, leader peptide, and presequence are used interchangeably and all refer to the peptide sequence found at the amino terminus of an encoded precursor protein that is subsequently removed to provide the secreted, mature form of the protein.

(c) Itakura *et al.*, "Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin," Science 198: 1056-1063 (1977) (Exhibit 2080)

(d) Villa-Komaroff *et al.*, "A bacterial clone synthesizing proinsulin," Proc. Nat'l Acad. Sci. USA 75: 3727-3731 (1978) (Exhibit 2081)

(e) Mercereau-Puijalon *et al.*, "Synthesis of an ovalbumin-like protein by *Escherichia coli* K12 harbouring a recombinant plasmid," Nature, 275: 505-510 (1978) (Exhibit 2082)

(f) Chang *et al.*, "Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dihydrofolate reductase," Nature 275: 617-624 (1978) (Exhibit 2083)

(g) Fraser and Bruce, "Chicken ovalbumin is synthesized and secreted by *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 75: 5936-5940 (1978) (Exhibit 2084)

(h) Seeburg *et al.*, "Synthesis of growth hormone by bacteria," Nature 276: 795-798 (1978) (Exhibit 2085)

(i) Goeddel *et al.*, "Expression in *Escherichia coli* of chemically synthesized genes for human insulin," Proc. Nat'l Acad. Sci. USA 76: 106-110 (1979) (Exhibit 2086)

(j) Schell *et al.*, "Cloning and expression of the yeast galactokinase gene in an *Escherichia coli* plasmid," Gene, 5: 291-303 (1979) (Exhibit 2087)

(k) Bach *et al.*, "Evidence for transcriptional regulation of orotidine-5'-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. 76: 386-390 (1979) (Exhibit 2088)

(l) Burrell *et al.*, "Expression in *Escherichia coli* of hepatitis B virus DNA sequences cloned in plasmid pBR322," Nature 279: 43-47 (1979) (Exhibit 2089)

(m) Martial *et al.*, "Human growth hormone: Complementary DNA cloning and expression in bacteria," Science 205: 602-607 (1979) (Exhibit 2090)

(n) Goeddel *et al.*, "Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone," Nature 281: 544-548 (1979) (Exhibit 1057)

(o) Wilson *et al.*, "Detection of proteins like human gamma and beta globins in *Escherichia coli* carrying recombinant DNA plasmids," Proc. Nat'l Acad. Sci. USA 76: 5631-5635 (1979) (Exhibit 2092)

(p) Roberts *et al.*, "Synthesis of simian virus 40 t antigen in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 76: 5596-5600 (1979) (Exhibit 2093)

(q) Roberts *et al.*, "A general method for maximizing the expression of a cloned gene," Proc. Nat'l Acad. Sci. USA 76: 760-764 (1979) (Exhibit 1126)

(r) Emtage *et al.*, "Influenza antigenic determinants are expressed from haemagglutinin genes cloned in *Escherichia coli*," Nature 283: 171-174 (1980) (Exhibit 2094)

(s) Steitz, "RNA·RNA Interactions During Polypeptide Chain Initiation," in RIBOSOMES Structure, Function and Genetics (Proceedings of the 9th. Steenbock Symposium Held at the University of Wisconsin, Madison, July 5-8, 1979; Ed. Chambliss *et al.*, University Park Press, Baltimore) (Exhibit 1109)

(t) Shine and Dalgarno, "Determination of cistron specificity in bacterial ribosomes," Nature 254: 34-38 (1975) (Exhibit 1043)

(u) Hautala *et al.*, "Increased expression of a eukaryotic gene in *Escherichia coli* through stabilization of its messenger RNA," Proc. Nat'l Acad. Sci. USA 76: 5574-5578 (1979) (Exhibit 1110)

(v) Goldberg, "Degradation of abnormal proteins in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 69: 422-426 (1972) (Exhibit 1111)

(w) Simon *et al.*, "Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 80: 2059-2062 (1983) (Exhibit 1112)

(x) Uren, "The recovery of genetically engineered mammalian proteins," Am. Biotechnol. Lab. 2: 51-54 (1983) (Exhibit 1113)

(y) Epstein, "The genetic control of tertiary protein structure: Studies with model systems," Cold Spring Harbor Symp. Quant. Biol. 28: 439-449 (1963) (Exhibit 1114)

(z) Fiers *et al.*, "Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene," Nature 260: 500-507 (1976) (Exhibit 1115)

(aa) Taniguchi *et al.*, "Construction and identification of a bacterial plasmid containing the human fibroblast interferon sequence," Proc. Jpn. Acad. 55: 464-469 (1979) (Exhibit 2059)

(bb) Maxam and Gilbert, "A new method for sequencing DNA," Proc. Nat'l Acad. Sci. USA 74: 560-564 (1977) (Exhibit 1116)

(cc) Goeddel, "Synthesis of human fibroblast interferon by *E. coli*," Nucl. Acids Res. 8: 4057-4074 (1980) (Exhibit 1012)

(dd) English Translation of Japanese Patent Application No. 033931/80 (Exhibit 2013)

(ee) Gray *et al.*, "Extracellular nucleases of Pseudomonas BAL 31. I. Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclease activities," Nucl. Acids Res. 2: 1459-1492 (1975) (Exhibit 1117)

(ff) Goeddel *et al.*, U.S. Patent No. 4,342,832, "Method of constructing a replicable cloning vehicle having quasi-synthetic genes," issued August 3, 1982 (Exhibit 1058)

(gg) Ptashne *et al.*, U.S. Patent No. 4,418,149, "Fused hybrid gene," issued November 29, 1983 (Exhibit 1118)

(hh) Ptashne *et al.*, U.S. Patent No. 4,332,892, "Protein synthesis," issued September 22, 1982 (Exhibit 2004)

(ii) Itoh *et al.*, "Efficient expression in *Escherichia coli* of a mature and a modified human interferon-β_1," DNA 3: 157-165 (1984) (Exhibit 1119)

(jj) UK Patent Application GB 2071671 (Exhibit 1078)

(kk) Itoh *et al.*, U.S. Patent No. 4,686,191, "Recombinant plasmid containing human interferon-beta gene," issued June 11, 1987 (Exhibit 1120)

(ll) Taniguchi *et al.*, "Expression of the Human Interferon-β_1 Gene in Heterologous Host Cells," *IN* "Proceedings of the IVth International Symposium on Genetics of Industrial Microorganisms (1982)" (Y. Ikeda & T. Beppu, eds.; Kobansha Ltd., Tokyo; 1983) (Exhibit 1068)

(mm) European Patent Application No. 0 064 681 A2 (Exhibit 1121)

(nn) Rosenberg and Court, "Regulatory sequences involved in the promotion and termination of RNA transcription," Ann. Rev. Genetics 13: 319-353 (1979) (Exhibit 1122)

(oo) Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009)

(pp) Goeddel *et al.*, "Human leukocyte interferon produced by *E. coli* is biologically active," *Nature* 287: 411-416 (1980) (Exhibit 2126)

(qq) Wei, "Isolation and comparisons of two molecular species of the BAL 31 nuclease from *Alteromonas espejiana* with distinct kinetic properties," J. Biol. Chem. 258: 13506-13512 (1983) (Exhibit 1123)

(rr) Backman *et al.*, "Maximizing gene expression on a plasmid using recombination *in vitro*," Cell 13: 65-71 (1978) (Exhibit 2095)

(ss) Declaration of Jordan U. Gutterman, M.D., dated January 24, 2007 (Exhibit 1026)

(tt) Taniguchi *et al.*, "The nucleotide sequence of human fibroblast interferon cDNA," Gene 10: 11-15 (1980) (Exhibit 1013)

(uu) Taniguchi *et al.*, "Molecular cloning of human fibroblast interferon cDNA," Proc. Nat'l Acad. Sci. USA 77: 4003-4006 (1980) (Exhibit 10146)

(vv) Chapter 5: "Antigen-Antibody Interactions," PRINCIPLES OF IMMUNOLOGY, 2nd Ed., MacMillan Publishing Co., New York (1979), pp. 65-79 (Exhibit 1016)

(ww) Sutcliffe, "Complete Nucleotide Sequence of the *Escherichia coli* Plasmid pBR322," Cold Spring Harb Symp Quant Biol., 43 Pt 1:77-90 (1979) (Exhibit 1032)

(xx) Knight, "Interferon: Purification and initial characterization from human diploid cells," Proc. Nat'l Acad. Sci. PNAS 73: 520-523 (1976) (Exhibit 1036)

(yy) Knight, "Human fibrboblast interferon: Amino acid sequence analysis and amino terminal amino acid sequence," Science 207: 525-526 (1980) (Exhibit 1037)

(zz) Houghton *et al.*, "The absence of introns within a human fibroblast interferon gene," Nucl. Acids Res. 9: 247-266 (1981) (Exhibit 1067)

(aaa) Jackson *et al.*, "Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 69: 2904-2909 (1972) (Exhibit 1124)

(bbb) Lobban and Kaiser, "Enzymatic end-to end joining of DNA molecules" J. Mol. Biol. 78: 453-471 (1973) (Exhibit 1125)

(ccc) Bendig *et al.*, "Deletion mutants of polyoma virus defining a nonessential region between the origin of replication and the initiation codon for early proteins," J. Virol. 32: 530-535 (1979) (Exhibit 1127)

(ddd) Shenk *et al.*, "Biochemical method for mapping mutational alterations in DNA with S1 nuclease: The location of deletions and temperature-sensitive mutations in simian virus 40," Proc. Nat'l Acad. Sci. USA 72: 989-993 (1975) (Exhibit 1128)

(eee) Grantham *et al.*, "Codon catalog usage and the genome hypothesis," Nucl. Acids Res. 8: r49-r62 (1980) (Exhibit 1129)

(fff) U.S. patent application Serial No. 06/131,152, which was filed March 17, 1980 (Exhibit 1130)

(ggg) Weissmann, "The Cloning of Interferon and Other Mistakes," IN INTERFERON 3, Academic Press Inc. (New York, NY), 1981, pp. 101-134 (Exhibit 1131)

(hhh) Curtiss, "Genetic manipulation of microorganisms: Potential benefits and biohazards," Ann. Rev. Microbiol. 30: 507-533 (1976) (Exhibit 2067)

(iii) Wu *et al.*, "Synthetic oligodeoxynucleotides for analyses of DNA structure and function," Prog. Nucl. Acid Res. Mol. Biol. 21: 101-141 (1978) (Exhibit 2068)

(jjj) Erlich *et al.*, "A sensitive radioimmunoassay for detecting products translated from cloned DNA fragments," Cell 13: 681-689 (1978) (Exhibit 2069)

(kkk) Guarente *et al.*, "Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin," Cell 20: 543-553 (1980) (Exhibit 2132)

9. In my opinion, a person skilled in the field of bacterial protein expression as of March 19, 1980 would have a Ph.D. degree, or be a highly skilled pre-doctoral fellow, with several years of research experience in molecular biology.

10. In preparing this Declaration, I have reviewed and considered the art cited above in light of my own, contemporaneous research in this field. The observations, opinions, and statements provided below reflect, in my view, the knowledge and understanding of those skilled in the art of heterologous expression of proteins in bacteria as of March 19, 1980.

11. It is my opinion that, as of March 19, 1980, one of ordinary skill in the art would not have had any reasonable expectation of either successfully constructing a DNA molecule encoding mature human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, or successfully expressing biologically active mature human fibroblast interferon in *E. coli*, without excessive experimentation. This was so even if

one of ordinary skill in the art had been provided with the nucleotide sequence of the coding region of the human fibroblast interferon precursor protein. That is, expression of mature human fibroblast interferon in *E. coli* as of March 19, 1980 was anything but "routine." These opinions are based upon my review of the art cited above, the evidence and reasoning provided below, and my own personal, contemporaneous attempts to express mammalian proteins, including mature human fibroblast interferon, in *E. coli* using, *inter alia*, enzymes and methods discussed below.

I. INTRODUCTION

12. As of March 1980, it was apparent that human fibroblast interferon had been identified as a potentially-useful human therapeutic agent for the treatment of viral infection and cancer. It was also apparent that the level and purity of human fibroblast interferon that could be provided by mammalian cell culture would not be sufficient to support clinical trials much less commercial-scale production.

13. At this time (March 19, 1980) there was a great deal of interest in the nascent field of genetic engineering, which was based, in part on the hope that the tools of recombinant DNA methodology might enable production of human therapeutic proteins in *E. coli*. This was particularly so in view of the fact that alternative sources of such proteins were extremely unattractive.

14. As demonstrated below, however, the art did not provide a reasonable expectation that, as of March 19, 1980, a DNA encoding mature human fibroblast interferon, unaccompanied by its signal peptide, could be constructed or that biologically active mature human fibroblast interferon, unaccompanied by the human fibroblast signal peptide, could be expressed in *E. coli*, without undue experimentation. Notwithstanding these observations, it was obvious to try to do so in view of the lack of alternative routes to obtain substantial quantities of mature human fibroblast interferon and in view of the potentially-tremendous clinical and commercial

importance of mature human fibroblast interferon. It was also obvious to try to express biologically active mature human fibroblast interferon in *E. coli* in light of the fact that success would likely provide enormous professional recognition. This organism, *E. coli*, was the host strain of choice at this time in light of the detailed genetics that had been developed for this organism, as well as its plasmids and phage, its non-pathogenic nature, and the existence of a method for transformation of this organism.

15. It should be noted that the field of recombinant DNA technology was expanding and advancing rapidly in the period spanning the years 1979 through 1985. Consequently, approaches that were in fact unreasonable or impossible in a practical sense in 1979-1980 were routine by 1985. Care must therefore be taken to view the art objectively and only from the vantage point one of ordinary skill as of, in the present instance, March 19, 1980. Thus, solutions that may be apparent in 2007 were not necessarily obvious or even available to those of ordinary skill faced with a specific problem in March 1980. In view of the state of the art as it existed prior to March 19, 1980, it is my opinion that those of ordinary skill did not have a reasonable expectation that a DNA encoding mature human fibroblast interferon, unaccompanied by its signal peptide, could be constructed and that biologically active mature human fibroblast interferon, having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, could be expressed in *E. coli.*

II. HETEROLOGOUS GENE EXPRESSION IN *E. COLI* AS OF MARCH 19, 1980

A. Introduction

16. Initial attempts to demonstrate expression of heterologous proteins in *E. coli* relied on genetic selection methods or on the construction of recombinant genes expressing fusion proteins, as will be discussed below. These initial disclosures did not provide a detailed teaching or even guidance concerning methods that could be used to tailor a mammalian gene so

that it could be used to express the encoded protein in a biologically active form in *E. coli*. Moreover, many of the reports available as of March 19, 1980, which are discussed below, related to the expression of cytoplasmic proteins. In view of the differences in structure, expression, and function that were known to exist between a cytoplasmic protein and a secreted mammalian glycoprotein, successful expression of the former was not considered predictive of success with the latter. However, as of March 19, 1980, there were two reports of mammalian proteins that had been directly expressed in *E. coli* (Roberts *et al.* (Exhibit 2093) and Goeddel *et al.* (Exhibit 1057)). Neither provided those of ordinary skill with a reasonable expectation that biologically active, mature human fibroblast interferon, having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, could be expressed in *E. coli* without undue experimentation as of March 19, 1980.

B. Heterologous expression by genetic selection

17. Publications available as of March 19, 1980, included reports of expression of a number of cytoplasmic proteins with a biochemical activity that could be selected for in *E. coli*. In this approach, a DNA molecule encoding a protein having the identified biochemical activity was introduced into an appropriate mutant *E. coli* host and transformants were selected by complementation of a genetic defect in the host strain. In some instances random DNA fragments of chromosomal DNA of a lower eukaryote were generated and inserted into an *E. coli* host strain. In such experiments, the power of the selection system allowed the isolation of clones in which the fortuitous combination of bacterial regulatory elements and eukaryotic gene resulted in sufficient expression to provide complementation of the genetic defect in the host cell (*see e.g.* Steitz (Exhibit 1109), at 490, last ¶, which ends on 491). In some instances, noted below, additional mutations and/or genetic rearrangements apparently were involved in facilitating expression of the gene from the lower eukaryote. In a similar manner, transformants

were also selected where the eukaryotic target protein provided resistance to an antibiotic to which the *E. coli* host strain was sensitive. In both instances, the ability to select transformants was taken as an indication that the heterologous gene had been expressed in *E. coli*. Although the methods described in these reports allowed, for example, the isolation of a specific eukaryotic gene, they did not provide those of ordinary skill in the art with any guidance regarding deliberate tailoring of a mammalian gene to enable heterologous expression of the encoded protein in an *E. coli* host. At most, these early publications did dispel concerns that eukaryotic coding sequences were inherently incapable of expression in *E. coli*. They did not, however, provide a detailed teaching of a method that could be applied to the expression of any specific mammalian protein in *E. coli*. Accordingly, as of March 19, 1980, these early references, even collectively, did not provide a basis for those skilled in the art to expect that the mature form of human fibroblast interferon, unaccompanied by the human fibroblast interferon signal peptide, could be expressed in *E. coli*.

18. Ratzkin and Carbon, "Functional expression of cloned yeast DNA in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 74: 487-491 (1977) (Exhibit 2078). Ratzkin and Carbon disclose the construction of a yeast (*Saccharomyces cerevisiae*) gene bank by cloning a set of randomly-generated yeast DNA fragments into an *E. coli* plasmid. The gene bank was introduced into an *E. coli* leucine auxotroph (*i.e.*, a mutant which requires nutritional supplementation with the amino acid leucine for growth) and colonies capable of growth on selective media, *i.e.*, without leucine, were selected. The restriction pattern of plasmid DNA from the fastest-growing clone was "quite different" from that of any other leu^+ clone. Exhibit 2078, at 489, left col., ll. 15-20. Ratzkin and Carbon did not disclose the mechanism underlying this heterologous expression of yeast DNA in *E. coli*, other than to suggest that "somehow"

barriers to transcription and translation had been overcome. Exhibit 2078, at 491, left col., third full ¶. The authors did not provide any further description of genetic changes or DNA rearrangements resulting in the observed heterologous expression of yeast DNA in *E. coli*.

19. Vapnek *et al.*, "Expression in *Escherichia coli* K-12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*," Proc. Nat'l Acad. Sci. USA 74: 3508-3512 (1977) (Exhibit 2079). Vapnek describes the construction of a *Neurospora* gene bank in *E. coli* in which DNA digested with the restriction enzymes EcoRI and HindIII was joined to similarly-digested pBR322 (plasmid vector) DNA. The gene bank was introduced into an *E. coli aro*D6 mutant, which is incapable of synthesizing the enzyme catabolic dehydroquinase, and transformed cells plated on selective media. One slowly-growing colony was identified and from that a "faster growing" colony was isolated and analyzed. Exhibit 2079, at 3510, left col., last full ¶. The authors noted that there appears to be a bacterial allele that determines whether or not the cloned DNA will complement the mutation in the bacterial host. Again, the authors did not provide any further description of genetic changes or DNA rearrangements leading to heterologous expression of this *Neurospora* gene in *E. coli*.

20. Chang *et al.*, "Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dihydrofolate reductase," Nature 275: 617-624 (1978) (Exhibit 2083). The murine dihydrofolate reductase (DHFR) gene had been cloned into the PstI site of the *E. coli* plasmid vector, pBR322, by GC tailing. Murine DHFR provides resistance to the drug trimethoprim while *thy*$^+$ *E. coli* strains are sensitive to trimethoprim (*tmr*). Accordingly, a population of *thy*$^+$ *E. coli* cells transformed with recombinant molecules carrying the murine DHFR gene were analyzed for those phenotypically expressing the murine gene by selecting for *tmr* resistant colonies. Chang suggested that expression of this gene might have been the fortuitous result of a

poly-G sequence, which resembled the Shine-Dalgarno sequence (AAGGAGGT), upstream from the ATG initiation codon for this cytoplasmic enzyme. However, no direct support for this inference was provided (*see* Roberts *et al.* (Exhibit 2093), at 5600, left col., ¶5, second sentence).

21. Schell *et al.*, "Cloning and expression of the yeast galactokinase gene in an *Escherichia coli* plasmid," Gene 5: 291-303 (1979) (Exhibit 2087). Schell described the cloning of a restriction fragment from yeast that, when inserted into the *E. coli* plasmid vector, pBR322, apparently resulted in the production of a *very* low level of galactokinase. That is, the enzymatic activity could be detected in a host *E. coli* carrying a deletion of the gene encoding bacterial galactokinase, effectively eliminating any background galactokinase activity. Schell did not describe the basis for expression of this heterologous yeast gene in *E. coli*.

22. Bach *et al.*, "Evidence for transcriptional regulation of orotidine-5'-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. 76: 386-390 (1979) (Exhibit 2088). The goal of this research apparently was the isolation of the yeast gene for orotidine-5'-phosphate decarboxylase, which was to be used as a probe to measure mRNA levels and regulation of transcription in *Saccharomyces cerevisiae* (Exhibit 2088, at 389, Fig. 4 and Table 2). Bach constructed a gene bank by inserting randomly-generated fragments of *Saccharomyces cerevisiae* DNA into the *E. coli* plasmid vector pMB9. Pooled plasmid DNA was used to transform an *E. coli* pyrimidine auxotroph host strain carrying *pyr*F mutation, which is a defect in orotidine-5'-phosphate decarboxylase, the last step in the biosynthesis of pyrimidines. Two clones were selected and characterized as carrying a yeast DNA fragment encoding and expressing orotidine-5'-phosphate decarboxylase. Although the authors speculated that the cloned DNA might carry a promoter

recognized by *E. coli*, no information was provided regarding the mechanism or structures responsible for the expression of yeast orotidine-5'-phosphate decarboxylase in *E. coli*.

C. Heterologous Expression by Gene Fusion

23. A second approach to demonstrating heterologous gene expression in *E. coli* relied upon the construction of gene fusions in which the coding sequence for the target protein was joined to a DNA molecule carrying a naturally-occurring, functional combination of promoter, ribosome binding site and amino-terminal coding region of an *E. coli* protein. The amino-terminal amino acids of a fusion protein encoded by such recombinant molecules were derived from an *E. coli* protein while the carboxy-terminal amino acids were derived from the heterologous target protein. However, those of ordinary skill in the art recognized that such fusion proteins would not be suitable as a human therapeutic agent. In this context, see Taniguchi *et al*., "Molecular cloning of human fibroblast interferon cDNA," Proc. Nat'l Acad. Sci. USA 77: 4003-4006 (1980) (Exhibit 1014), at page 4005: "[i]t will be quite important to let the gene express correctly, because fused or incomplete proteins synthesized in bacteria may not be useful in clinical applications, even if such proteins exhibit antiviral properties." It is also my understanding that this issue has been addressed, in detail, by a Declaration from Jordan U. Gutterman, M.D. (Exhibit 1026) submitted previously in these proceedings.

24. Gene fusions in which the amino-terminal coding region was derived from a normally secreted *E. coli* protein (*e.g.*, β-lactamase), were used to investigate the possibility that the fusion protein might be transported across the cell membrane to the periplasmic space of *E. coli*. In addition, gene fusions encoding heterologously expressed peptides joined to large prokaryotic proteins, were constructed in an attempt to protect therapeutically useful peptides from proteolysis. Synthesis of an apparently correctly folded protein or at least an epitope thereof was frequently only established using extremely sensitive immunoassays ("RIA").

Again, such reports did not provide those of ordinary skill with a detailed teaching of a method that could be applied to the expression of any specific mammalian protein in *E. coli*. Moreover, even this art included apparently contradictory results regarding transport of fusion proteins to the *E. coli* periplasmic space. That is, hybrid proteins including the amino-terminal portion of β-lactamase fused to proinsulin or fused to ovalbumin were reported as either capable of being transported to the periplasmic space of (Villa-Kamaroff *et al.* (Exhibit 2081); Fraser and Bruce (Exhibit 2084)) or as retained within (Seeburg *et al.* (Exhibit 2085); Mercereau-Puijalon *et al.* (Exhibit 2082)) the *E. coli* host cell. These data illustrated that whether or not the fusion protein could be transported into the periplasmic space was polypeptide-specific.

25. Itakura *et al.*, "Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin," Science 198: 1056-1063 (1977) (Exhibit 2080). Itakura disclosed the chemical synthesis of DNA encoding the 14-amino acid peptide hormone, somatostatin, selecting, where experimentally feasible, codons commonly used ("preferred") by the intended host, *E. coli*. Attempts to express somatostatin fused to a bacterial peptide sequence of comparable length did not yield any detectable product, apparently because this relatively short fusion peptide of fewer than 30 amino acids was proteolytically destroyed within *E. coli*. Accordingly, the synthetic DNA was inserted at the carboxy-terminal end of the coding sequence for the bacterial protein, β-galactosidase. Intracellular expression of this hybrid gene provided a fusion protein from which somatostatin could be isolated by chemical digestion with cyanogen bromide.

26. Villa-Komaroff *et al.*, "A bacterial clone synthesizing proinsulin," Proc. Nat'l Acad. Sci. 75: 3727-3731 (1978) (Exhibit 2081). Villa-Komaroff inserted a rat proinsulin cDNA into the PstI site of pBR322 by GC tailing procedures. The PstI site is within the coding

sequence for the normally secreted, periplasmic enzyme, β-lactamase (ampicillin resistance). One clone was identified as encoding an in-frame fusion protein of 182 amino acids of β-lactamase, followed by 6 glycine residues (GGG encodes glycine) and all but the first 4 amino acids of rat proinsulin. Osmotic shock provided material detectable by an extremely sensitive RIA method. This observation was interpreted as suggesting that *at least some* of the fusion protein comprising rat proinsulin sequences had been transported through the cell membrane and that a proinsulin epitope could be recognized by an antibody.

27. Mercereau-Puijalon *et al.*, "Synthesis of an ovalbumin like protein by *Escherichia coli* K12 harbouring a recombinant plasmid," Nature 275: 505-510 (1978) (Exhibit 2082). Mercereau-Puijalon discloses the construction and expression of a fusion protein consisting of 392 amino acids. The amino-terminal 8 amino acids are derived from *E. coli* β-galactosidase and the remaining 384 amino acids from the normally-secreted chicken ovalbumin protein. Transcription of this fusion gene was regulated by the *lac* promoter and the translation initiation signals were those of β-galactosidase. Although substantial amounts of the β-galactosidase-ovalbumin fusion protein accumulated within the cell, this fusion protein apparently was not transported across the bacterial membrane. RIA competition experiments (reported at Exhibit 2082, at 507) suggested that one or more antigenic sites were missing from the bacterially produced ovalbumin sequences.

28. Fraser and Bruce, "Chicken ovalbumin is synthesized and secreted by *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 75: 5936-5940 (1978) (Exhibit 2084). Fraser and Bruce described the construction of a hybrid gene encoding a fusion protein consisting of 7 amino-terminal amino acids of *E. coli* β-galactosidase, 11 amino acids encoded by DNA upstream of the ATG initiation codon of ovalbumin, followed by the coding region for chicken

ovalbumin. The authors indicated that the amount of fusion protein was about 10% of the amount predicted, and ascribed this discrepancy to a lower rate of translation of the heterologous chicken sequences in *E. coli*. The authors suggested that the low level of ovalbumin synthesis observed could be the result of the different codon frequency usage between *E. coli* proteins and that present in the mammalian, ovalbumin coding sequence (Exhibit 2084, at 5940, left col., first four lines). Fraser and Bruce further reported that approximately 50% of the β-galactosidase-ovalbumin fusion protein was found in the *E. coli* periplasmic space. These observations are in contrast to the report of Mercereau-Puijalon *et al.* (Exhibit 2082), which indicated that a structurally similar fusion protein, also expressed in *E. coli* was not transported through the bacterial membrane.

29. Seeburg *et al.*, "Synthesis of growth hormone by bacteria," Nature 276: 795-798 (1978) (Exhibit 2085). Seeburg reported the construction of a hybrid gene encoding a fusion protein that included the amino-terminal 182 amino acids of bacterial β-lactamase joined, in frame, to the coding sequence of the rat proinsulin gene, including 24 amino acids of the proinsulin leader or signal peptide. Seeburg reported that this fusion protein could be detected by RIA. However, in contrast to the results reported by Villa-Komaroff *et al.* (Exhibit 2081), the β-lactamase-rat proinsulin fusion of Seeburg apparently was not found in the periplasmic space.

30. Goeddel *et al.*, "Expression in *Escherichia coli* of chemically synthesized genes for human insulin," Proc. Nat'l Acad. Sci. USA 76: 106-110 (1979) (Exhibit 2086). Goeddel reports the chemical synthesis and enzymatic assembly of separate DNA molecules encoding the A and B chains of human insulin. Each such DNA was separately fused to the coding sequence of *E. coli* β-galactosidase. Transcription and translation of each hybrid gene were regulated by the upstream *lac* promoter and SD sequence associated with β-galactosidase. The fusion proteins

were synthesized at high levels and found to be insoluble. The A and B insulin chains were released from the fusion proteins by treatment with cyanogen bromide, oxidized and then assembled *in vitro* to provide authentic human insulin. As noted at page 106, this approach was facilitated by ongoing developments in the field of chemical synthesis of oligodeoxyribonucleotides.

31. Burrell *et al.*, "Expression in *Escherichia coli* of hepatitis B virus DNA sequences cloned in plasmid pBR322," Nature 279: 43-47 (1979) (Exhibit 2089). Burrell disclosed the insertion of HBV DNA restriction fragments into the PstI site of pBR322 using GC tailing methods. Although a number of the hybrid plasmids were apparently unstable, Burrell did identify, by RIA, thirteen clones as carrying HBcAg. The clones apparently expressed a protein in which HBcAg sequences had been fused to the amino-terminal portion of plasmid-encoded β-lactamase. Burrell suggested that some of this fusion protein was transported to the periplasmic space.

32. Martial *et al.*, "Human growth hormone: Complementary DNA cloning and expression in bacteria," Science 205: 602-607 (1979) (Exhibit 2090). Martial disclosed construction and cloning of the cDNA for human growth hormone as well as the DNA sequence encoding 191 amino acids of the mature protein as well as the 26 amino acids of the signal peptide. Martial also disclosed expression of hGH as a fusion protein comprising the amino-terminal coding region of the *E. coli trp*D protein, amino acids encoded by the nucleotide sequence upstream of the hGH gene, the 26-amino acid signal peptide, and, finally, the 191 amino acids of the mature form of hGH. This complex fusion protein could, nonetheless, be precipitated by antiserum to hGH. The authors noted that, although this fusion protein does not appear to be unstable, it only accumulated at 17% of the expected level. By way of explanation,

the authors referred to the Fraser and Bruce article noted above (Exhibit 2084), in which a slower rate of translation (possibly the result in codon frequency differences) was suggested as responsible for expression levels lower than those predicted.

33. Wilson *et al.*, "Detection of proteins like human γ and β globins in *Escherichia coli* carrying recombinant DNA plasmids," Proc. Nat'l Acad. Sci. USA 76: 5631-5635 (1979) (Exhibit 2092). Wilson disclosed expression of RIA-detectable human β-globin and γ-globin sequences in *E. coli*. Wilson noted that their data regarding antibody inhibition (competition) assays indicated that their bacterially synthesized protein may have an altered conformation or may even be a fusion protein comprising bacterial sequences. Wilson also suggested that "bacterial mechanisms" were responsible for transcription of the globin genes. Exhibit 2092, at 5633, left col., third full ¶ and at 5635, left col., second full ¶.

34. Emtage *et al.*, "Influenza antigenic determinants are expressed from haemagglutinin genes cloned in *Escherichia coli*," Nature 283: 171-174 (1980) (Exhibit 2094). Emtage discloses the cloning of a DNA fragment encoding fowl plague virus (FPV) hemagglutinin (HA) into the HindIII site of the *E. coli* plasmid cloning vector. This publication described the construction of an expression plasmid carrying an *E. coli trp* promoter. In one construction the expected fusion protein would encompass: 7 amino-terminal amino acids of antranilate synthetase (*E. coli trp* enzyme), 6 amino acids encoded by the linker DNA, 6 phenylalanine residues encoded by the $(T)_{19}$ region of the FPV cloned DNA, 7 amino acids from the normally-untranslated upstream sequences of the HA gene, 558 amino acids of FPB HA including the HA signal peptide, and 5 amino acids encoded by the linker at the carboxy-terminus. This complex fusion protein has a calculated molecular weight of 69 kD. Expression of this fusion protein, as well as the protein encoded by two other recombinant

molecules in which the HA coding sequence was in a reading frame different from that of *trp* coding sequence, all generated RIA-detectable material (see Exhibit 2094, at 173, Table 1). However, in each instance, the apparent molecular weight of the RIA-detectable product was found to be 61 kD (Exhibit 2094, at 174, left col., first sentence). Although the authors ascribed this discrepancy to proteolytic processing, these data can also be explained by initiation of translation from an internal, in-frame ATG (methionine codon) yielding a truncated HA polypeptide that could be recognized by the anti FPV-HA serum used in the RIA. Such events have been documented *e.g.*, by Roberts *et al.* (Exhibit 2093) and Taniguchi *et al.* (Exhibit 1009). Thus it is not at all clear that translation of the coding sequence for HA protein was initiated with the ATG codon recognized *in vivo* (*i.e.*, in a mammalian cell).

35. The construction of gene fusions expressing hybrid proteins precluded the need for tailoring the heterologous gene to be expressed by exploiting naturally-occurring prokaryotic genetic elements known to be effective for expression of the *E. coli* carrier protein. Accordingly, it is my opinion that these reports did not provide those of ordinary skill in the art with guidance or a detailed teaching of a method for deliberate tailoring of a mammalian gene to enable heterologous expression in an *E. coli* host of an encoded target mammalian protein, such as biologically active mature human fibroblast interferon, having a total of 165 or 166 amino acids, and unaccompanied by the human fibroblast interferon signal peptide.

D. Expression of a Cytoplasmic Protein as Compared to Expression of a Secreted Mammalian Glycoprotein

36. As noted above, those of ordinary skill were aware of the chemical and biological differences between an intracellular, cytoplasmic protein and a secreted mammalian glycoprotein, and understood that success in expression of the former was not predictive of success in the expression of the latter. This inference of those of ordinary skill in the art was

based upon a number of observations. For example, the interior of the *E. coli* cell provides a reducing environment in which disulfide bonds are not formed between cysteine residues of a cytoplasmic protein. In contrast, secreted proteins are found in an extracellular, oxidizing environment in which disulfide bonds do form between cysteine residues of the protein and are expected to stabilize the properly folded, biologically active structure of the protein. In addition, since *E. coli* was known not be able to glycosylate proteins faithfully, those of ordinary skill were concerned that a nonglycosylated, heterologously expressed protein might not fold properly, particularly in the intracellular, reducing environment of an *E. coli* cell, and, therefore, might not be biologically active. Finally, secreted mammalian proteins were known to be synthesized as precursor proteins carrying an amino-terminal presequence (referred to herein as a "signal peptide") that is normally removed during the processes of synthesis, transport, and glycosylation of the protein. Moreover, since efficient and faithful systems for secretion of mammalian proteins expressed in *E. coli* were not available before March 19, 1980, those of ordinary skill in the art understood that one or more of these elements, either alone or in combination, could prevent the cytoplasmic production of a biologically active nonglycosylated mature form of a normally secreted mammalian glycoprotein, such as mature human fibroblast interferon, in *E. coli*. Thus, even though the mature form of human growth hormone had been expressed in *E. coli*, those of ordinary skill in the art still could not reasonably expect that another *normally glycosylated* mammalian protein could also be expressed in *E. coli* in biologically active form, especially since human growth hormone is not a glycoprotein.

37. Those of ordinary skill in the art as of March 1980 also knew that human fibroblast interferon was a very hydrophobic molecule and that its amino acid sequence included an odd number (three) of cysteine residues. In view of this hydrophobicity, and particularly in

the absence of glycosylation, there was a real concern that mature human fibroblast interferon protein would stick to cellular membranes and be toxic to the *E. coli* host cell in which it was expressed. For the same reasons, there was also a concern that nonglycosylated mature human fibroblast interferon would be poorly water soluble, creating obstacles to the expression, isolation, and biological assay of that protein. In addition, the presence of three cysteine residues in the mature form of fibroblast interferon was also potentially problematic in that formation of one intramolecular disulfide bond in an oxidizing environment would leave a third, unpaired sulfhydryl moiety. The latter moiety was expected to facilitate intermolecular cross-linking as well as the formation of incorrect intramolecular disulfide bonds resulting in the accumulation of aberrant protein structures and loss of biological activity. In contrast, the mature form of human growth hormone includes four cysteine residues that form two disulfide bonds.

38. The art available as of March 19, 1980 also included a report of the heterologous expression of a nonglycosylated, cytoplasmic viral protein, the small t antigen of SV40 in *E. coli* (Roberts *et al.* (Exhibit 2093)). In view of the absence of a signal peptide and the presence of fortuitously positioned restriction site immediately adjacent to the ATG initiation codon, manipulation of the gene encoding this protein for expression in *E. coli* could be relatively straightforward. This report demonstrated that a cytoplasmic, nonglycosylated mammalian protein could be expressed without fusion to an *E coli* peptide sequence. Those of ordinary skill in the art as of March 19, 1980, however, would have recognized that the methods reported would not be directly applicable, *e.g.*, to the tailoring of a DNA encoding the normally secreted and glycosylated, human fibroblast precursor protein, which included, *inter alia*, the coding sequence for a 21-amino acid signal peptide that would have to be removed. Accordingly, those of ordinary skill in the art did not reasonably expect that biologically active mature human

fibroblast interferon, having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, would be expressed in *E. coli*, based upon the data reported for t antigen, even in combination with the art.

39. Goeddel *et al*., "Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone," Nature 281: 544-548 (1979) (Exhibit 1057) ("Goeddel hGH paper"). In contrast to fibroblast interferon, hGH is not a glycoprotein, obviating at least one potential impediment to heterologous expression in *E. coli*. Moreover, since mature hGH includes an even number of cysteine residues (four) which form two disulfide bonds in an oxidizing environment, hGH does not carry a free cysteine residue. Expression of hGH, therefore also avoids the potential impediments to correct folding that arise from the presence of a free cysteine residue, *i.e.*, the formation of intermolecular disulfide bonds as well as incorrect, intramolecular disulfide bonds. However, direct expression of the mature form of human growth hormone ("hGH") in bacteria was particularly difficult in that the amino-terminal amino acid of the mature protein is phenylalanine. Therefore, expression of hGH in bacteria required not only removal of the coding sequence for the 26-amino acid hGH leader and proper positioning of a promoter and SD sequence, but also the creation of an initiation codon (ATG) immediately before the phenylalanine codon. In this instance, the first two amino acids of the expressed protein would be Met-Phe, and it was hoped that the *E. coli* host would remove the amino-terminal Met residue. Goeddel described the assembly of an expression system that relied on a complex series of plasmid constructions as well as the use of chemically synthesized oligonucleotides encoding the noted ATG codon as well as the first 24 amino acids of hGH (Fig. 1, p. 545). Goeddel also reported that the level of expression of hGH decreased almost three-fold when the spacing between the *lac* SD and the ATG initiation codon was decreased

from 11 base pairs to the 7 base pair spacing found in the *lac* operon (p. 546, left col., which ends at p. 547).. This report, therefore, also demonstrated that a mammalian protein could be expressed without fusion to an *E coli* peptide sequence. Those of ordinary skill in the art as of March 19, 1980, however, recognized that the methods reported would not be directly applicable, *i.e.*, did not enable, tailoring of a DNA encoding mature human fibroblast precursor protein. One reason for this was the absence of a unique restriction site near the ATG codon for the amino-terminal methionine residue of mature human fibroblast interferon, in contrast to the fortuitously occurring HaeIII restriction site exploited by Goeddel for expression of human growth hormone (*see* Goeddel *et al.* (Exhibit 1057), at 544, left col., third full ¶). For this reason, as well as for the reasons provided in the following paragraphs, the expression of human growth hormone in *E. coli* by Goeddel *et al.* (Exhibit 1057), did not provide those of ordinary skill in the art with a basis for a reasonable expectation that biologically active mature human fibroblast interferon, having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, would be expressed in *E. coli* without undue experimentation.

40. More specifically, the Goeddel hGH paper, described the first instance in which a normally secreted protein, which is encoded as a precursor protein, was directly expressed as a mature protein, without its signal peptide, in *E. coli.* This publication describes an approach of constructing a particular DNA encoding a mature protein, human growth hormone, for direct expression. To utilize this method, one must have a DNA encoding that portion of the precursor form of the protein including all of the mature form, and know the DNA sequence of that part of such DNA encoding the amino terminus of the mature portion of the protein to be expressed.

41. I was aware of the Goeddel hGH paper shortly after it was published.

42. The Goeddel hGH paper describes and exemplifies the approach for one particular protein, human growth hormone. The cDNA encoding human growth hormone was susceptible to cleavage with a single restriction endonuclease (HaeIII) at two sites such that a cleavage was made (at codon 24) close to the 3' side of the codon for the amino-terminal amino acid of the mature protein and a cleavage was made following the termination codon in the 3' untranslated region. The resulting HaeIII fragment included the coding sequence for amino acids 24-191 of human growth hormone. Then, a chemically synthesized DNA fragment representing the translational start signal (ATG) and coding sequences for amino acids 1-23 of human growth hormone (up to the cleavage site within the coding sequence) was prepared. These two DNA fragments were combined to form a synthetic-natural "hybrid" gene which was then inserted into a vector and positioned for proper expression of the mature human growth hormone protein (Fig. 4, p. 547).

43. Thus, as described in the Goeddel hGH paper, "mature" human growth hormone is directly expressed from a DNA sequence having a translational start signal (ATG) immediately preceding the codon corresponding to the first amino acid (*i.e.*, amino-terminal) of the mature protein encoded within the human gene. For mature human growth hormone, that amino-terminal amino acid is a phenylalanine residue. The ultimate protein resulting from direct expression will contain a methionine residue (resulting from the ATG translational start signal) as its amino-terminal amino acid unless the microbial host processes the protein to delete that initial methionine residue.

44. To the best of my knowledge, there were no published reports as of March 19, 1980 in which the approach proposed in the Goeddel hGH paper, or a variant thereof, had been used to directly express the mature form of any protein other than human growth hormone.

45. I am aware that near the end of the Goeddel hGH paper (Exhibit 1057, at 548, right col., third full ¶), the following statement appears:

> The hybrid DNA cloning techniques described as a route to the cloning and expression of HGH coding sequences in E. coli are generally applicable to other polypeptides which are synthesized initially as inactive precursors and later processed, or for which full length cDNA transcripts are unavailable.

Subsequent experience has confirmed that the described approach is, in fact, generally applicable. However, prior to March 19, 1980, and at least for the subsequent five years, the information provided in that paper did not allow one skilled in the art to design with a reasonable expectation of success a method that would be successful when applied in an attempt to produce a specific mammalian protein.

46. During the period of 1981-1987 when I was a researcher at Genentech under the supervision of Dr. Goeddel, the approach proposed in the Goeddel hGH paper was considered as a general directive to obtain expression of a mature protein in *E. coli*, but was not considered to predict successful bacterial expression of any specific mature recombinant protein. Furthermore, even for the in-house Genentech scientists with scientific skill beyond the ordinary skill in the art, the Goeddel hGH paper was not sufficiently instructive or enabling to be considered a method adaptable to the expression of other mature proteins in *E. coli*, without extensive modification and experimentation.

47. Before the generality of the approach of the Goeddel hGH paper was confirmed by its successful application to produce other biologically active mammalian proteins in addition to human growth hormone, it is my opinion that one skilled in the art could not properly or reasonably have concluded with any degree of reasonable expectation that the method would successfully work for any specific mammalian protein.

48. During 1979 and through March 19, 1980, I was aware that several research groups, including the research group of Dr. Walter Fiers, of which I was a member, were engaged in projects aimed at expressing various mammalian proteins in bacteria. In particular, the Fiers group and another group, the Tanaguchi/Ptashne group, were interested in expressing mature human fibroblast interferon during 1980. These groups had available to them the published Goeddel hGH paper. To my knowledge, none of these researchers predicted that the approach used in the Goeddel hGH paper could be generally successfully used to express in bacteria any specific mature mammalian protein.

49. The research reported in Goeddel *et al.*, "Human leukocyte interferon produced by *E. coli* is biologically active," Nature 287: 411-416 (October 2, 1980) (Exhibit 2126), disclosed that a variant of the method of the Goeddel hGH paper was successfully used to produce a mature human leukocyte interferon in biologically active form. Since human leukocyte interferon is not glycosylated, its expression would not have resolved the issues raised concerning expression of normally glycosylated human fibroblast interferon. To my knowledge, this was only the second publication disclosing that a biologically active mature mammalian protein was expressed in bacteria using elements of the approach used to express human growth hormone. Note, however, that this paper was published after March 19, 1980 and, therefore, was not available to the skilled researcher as of the filing date of the Japan '931 application.

E. METHODS AND ASSAYS

50. Erlich *et al.*, "A sensitive radioimmunoassay for detecting products translated from cloned DNA fragments," Cell 13: 681-689 (1978) (Exhibit 2069). Erlich discloses an immunoassay in which $F(ab)'_2$ fragments specific for a target protein, which are generated by pepsin digestion of antiserum, are attached to a solid support. The support-bound $F(ab)'_2$ are contacted with a test sample to capture the target protein. The presence of bound target is

demonstrated by further incubation with antiserum and ^{125}I-labeled *S. aureus* protein A. This method is very sensitive and can detect low levels of expression of a target protein or epitopes thereof. Accordingly, it is useful for the initial detection of clones carrying part or all of the coding sequence of a target gene product. However, an antibody raised against the mature form of a normally secreted protein (*e.g.*, mature human fibroblast interferon) would not be expected to be able to distinguish the mature form of that protein from the precursor form thereof, carrying a signal peptide sequence (or portion thereof). Moreover, as I recall, an RIA had not been established or reported for human fibroblast interferon as of March 19, 1980, despite the intense interest in this molecule exhibited by the Ptashne, Fiers, Goeddel, and Sugano laboratories.

51. Curtiss, "Genetic manipulation of microorganisms: Potential benefits and biohazards," Ann. Rev. Microbiol. 30: 507-533 (1976) (Exhibit 2067). Curtiss provides an overview of the potential benefits of the newly developed methods for manipulation of microorganisms using the tools of genetic engineering. Curtiss focuses on potential dangers arising from the release of genetically engineered microorganisms into the environment, and describes not only methods intended to prevent such release but also approaches that would minimize the survival of such organisms in a natural environment. As an introduction to this discussion, Curtiss merely provides a one paragraph summary entitled "Current Status of Technology," that begins at page 508 and ends on page 509. In this paragraph, Curtiss describes one of the earliest methods used for joining two DNA molecules to form a "recombinant DNA" (Jackson *et al.*, "Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 69: 2904-2909 (1972) (Exhibit 1124); and Lobban and Kaiser, "Enzymatic end-to end joining of DNA molecules" J. Mol. Biol. 78:

453-471 (1973) (Exhibit 1125)). More specifically, Curtiss refers to an approach in which a plasmid DNA (vector) and target (passenger) DNA molecules are first converted to linear molecules (*e.g.*, by digestion with a restriction endonuclease (EcoRI), and any "staggered" ends having protruding 5'-ends and recessed 3'-ends are treated with λ-exonuclease. This step removes protruding 5'-ends since λ-exonuclease digests DNA in the "5'-to-3' direction, thereby converting the recessed 3'-ends to protruding 3'-ends. That is, there is no precision required in this step beyond removing the single-stranded 5'-end. The resulting protruding 3'-ends, which are better substrates for the enzyme terminal transferase, are then extended by attachment of a homopolymeric tail of deoxyribonucleotide residues. In this manner, complementary tails are added to the vector (*e.g.*, a poly A tail) and the passenger (*e.g.*, a poly T tail) allowing the two molecules to be annealed by Watson-Crick base pairing. The four gaps created by this procedure are "filled in" with DNA polymerase I, and the strands covalently sealed with *E. coli* DNA ligase. These molecules may also be treated with the Exonuclease III ("Exo III") since ExoIII "removes 3'-phosphoryl residues at any nicks inadvertently introduced during the manipulations ("nicks with 3'-phosphoryl ends cannot be sealed by ligase") (*see* Jackson *et al.* (Exhibit 1124), at 2906, left col. first ¶, last sentence; *also see* Lobban and Kaiser (Exhibit 1125), at 462, last ¶, last 3 lines)). That is, Exo III was not used to remove any base pairs from either the vector or the passenger DNA molecules in the publications cited by Curtiss. In summary, the methods referred to by Curtiss are, at best, only remotely related to heterologous gene expression.

52. Wu *et al.*, "Synthetic oligodeoxynucleotides for analyses of DNA structure and function," Prog. Nucl. Acid Res. Mol. Biol. 21: 101-141 (1978) (Exhibit 2068). Wu provides a summary of the history of oligodeoxyribonucleotide synthesis. In particular, Wu notes that the synthesis of a 77-base pair duplex DNA gene encoding yeast alanine tRNA involved five years

of intensive work (Exhibit 2068, at 121, first full ¶, first sentence). Subsequent work in the same laboratory led to the synthesis, cloning, and expression of a second tRNA gene, a 207 base pair molecule that not only encoded the *E. coli* tyrosine suIII$^+$ tRNA (*supF*) precursor, but also included the naturally occurring promoter for that tRNA precursor. In addition, Wu provides a cursory review of ongoing developments related to the chemical synthesis of oligonucleotides. Wu further discusses the potential uses that could be made of synthetic oligonucleotides as probes, primers, protein binding substrates, and linkers for introduction of restriction sites. In particular, Wu notes that S1 nuclease could be used for removal of single-stranded, 3'-protruding ends. However, Wu warns that if the substrate DNA carries a single-stranded break, S1 may cleave the second strand of the duplex at that point, creating a double-stranded break; *i.e.*, cleaving the duplex into two fragments (Exhibit 2068, at 136, first ¶).

53. It is my opinion that the published reports discussed in ¶¶ 16-52 above, did not provide those of ordinary skill with a basis for a reasonable expectation that any mammalian protein of interest could be expressed in *E. coli*. The observation, *e.g.*, that a particular protein is stable in *E. coli* or that a particular protein can be expressed in *E. coli* was not predictive of successful expression of another *selected* protein. The reports discussed in ¶¶ 16-52 above therefore would have been considered by those of ordinary skill to be more of a collection of anecdotal reports than a coherent body of art. Thus, as of March 19, 1980, the references discussed above did not provide a basis for those skilled in the art to expect that biologically active, mature human fibroblast interferon, having a total of 165 or 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide, could be expressed in *E. coli*, without undue experimentation. It is also my opinion that this would be so even if one were provided with the coding sequence for the precursor form of human fibroblast interferon.

III. EXPRESSION OF HETEROLOGOUS TAILORED GENES IN *E. COLI*

A. Requirements for Transcription and Translation Initiation

54. As of March 19, 1980, it was known that expression of a gene in *E. coli* depended on the presence of an upstream promoter sequence that would be recognized by RNA polymerase for initiation of transcription of the gene to provide the corresponding messenger RNA (mRNA). Expression also required the presence of a "ribosome binding site" embedded within that mRNA for initiation of translation of the mRNA into protein. A sufficient number of *E. coli* promoters had been identified and analyzed by 1980 to allow the identification of "consensus sequence" of nucleotides recognized by *E. coli* RNA polymerase (Rosenberg and Court (Exhibit 1122)). In a similar manner, sequence analysis of the non-coding region upstream of the ATG initiation codon of a number of prokaryotic genes allowed the identification of a conserved sequence of nucleotides, designated the "Shine-Dalgarno" sequence ("SD sequence"), that appeared necessary for efficient initiation of translation. The phrase "ribosome binding site" refers to that portion of the mRNA including the SD sequence and the ATG initiation codon as well as the nucleotides disposed between the SD and the ATG. As of March 1980, it was recognized that the presence and relative disposition of a promoter, SD sequence and ATG were necessary, but not sufficient, for efficient expression of a gene in *E. coli*.

55. More specifically, accurate initiation of mRNA translation depends upon two sequences embedded within the mRNA. The first was the initiation codon ATG (which codes for methionine) and the second was designated a "Shine-Dalgarno" ("SD") sequence (*e.g.*, AAGGAGGU), which is complementary to the 3'–end of 16 S RNA of the 30 S ribosomal subunit of *E. coli*. The SD sequence ranged from 3 to 9 nucleotides in length and was centered, on average, approximately 10 nucleotides upstream from the ATG initiation codon (Steitz (Exhibit 1109), at 482). However, although the presence of a ribosome binding site, including a

SD sequence and ATG initiation codon, was *necessary* for translation initiation, it was not *sufficient* for translation initiation. That is, many ATG codons had been identified that were "appropriately juxtaposed" to a SD sequence, but were not used for translation initiation (Steitz (Exhibit 1109), at 491). It had also been recognized that translation initiation required the ATG codon to be accessible, *i.e.*, not embedded within a stable mRNA secondary or tertiary structure (Steitz (Exhibit 1109), at 483 and 486). Expression of a eukaryotic gene in *E. coli* was found to be "appallingly low" unless deliberately fused to an *E. coli* ribosome binding site (Steitz (Exhibit 1109), at 491). Finally, it had been noted that changing the distance between an intact ribosome binding site and the non-coding, upstream 5'-end of the mRNA, even by only a few nucleotides, could have a drastic effect on the level of expression of the encoded protein. It was proposed that the latter observation could be a reflection of mRNA stability and/or mRNA secondary and tertiary structure (Steitz (Exhibit 1109), at 492).

56. As of March 19, 1980, it was also understood that the spacing between the SD sequence and the ATG *within a specific mRNA* could have a dramatic effect on translation of that mRNA. Moreover, it was also apparent that the length, the nucleotide sequence, and therefore the secondary and tertiary structure, of *each specific mRNA*, including the non-coding sequences of the mRNA found "upstream" of the SD sequence and the ATG, could also have a dramatic effect on translation of that mRNA. For example, if the ATG and/or the SD sequences were to be sequestered within stable duplex structures (*e.g.*, embedded within the stem of a "hairpin" structure), they could be inaccessible to ribosomes. Consequently, initiation of translation using that ATG as an initiation codon could be inhibited or prevented. Moreover, in some instances, it was also possible that a "downstream" ATG could be used for translation initiation, yielding a completely different protein or a truncated version of the target protein, depending on the reading

frame of this "downstream ATG" (Roberts *et al.* (Exhibit 2093), at 5599, right col., first ¶ , last sentence; *also see* ¶ 34 above).

57. Therefore, as of March 19, 1980, those of ordinary skill understood that, as a general rule, expression of a gene encoding a protein required a promoter, SD sequence and an initiation codon. Moreover, those of ordinary skill in the art could, with a reasonable degree of certainly, infer that some combinations of promoter, SD sequence and ATG would not result in expression of the target protein; *e.g.*, one in which the SD sequence and ATG were separated by more than 25 base pairs (Steitz (Ex. 1109) at 480, first ¶, second sentence and at 481-482, Table 1). However, one of ordinary skill could not, *a priori*, state whether any particular combination of mRNA sequence (which of course differs for each protein), SD sequence and ATG, would result in expression of the gene in question.

58. In view of the information summarized in the preceding paragraphs, those of ordinary skill in the art as of March 19, 1980 recognized, if only in concept, that expression of a mammalian gene encoding a target protein required "tailoring" of that gene to provide a DNA molecule in which (a) an ATG initiation codon was placed immediately before the codon for the first amino acid of the target protein and (b) both a bacterial promoter and SD sequence were "appropriately" disposed upstream of that ATG initiation codon.

59. In view of the above, it is my opinion that a DNA encoding the mature form of human fibroblast interferon would be one that would be directly useful for its intended purpose, *i.e.*, for direct expression of mature human fibroblast interferon having a total of 166 amino acids and unaccompanied by its presequence. In view of the requirement that the SD sequence be closely associated with the ATG initiation codon, a DNA encoding the mature form of human fibroblast interferon would be one from which essentially the entire coding region for the signal

peptide had been removed (*see e.g.*, Goeddel *et al.* (Exhibit 1012), at 4066, Fig. 4 for illustrative examples of such DNA molecules encoding the mature form of human fibroblast interferon)). Thus, simply removing the ATG initiation codon from the coding sequence for the 187-amino acid precursor form of human fibroblast interferon would not, in my opinion, provide a coding sequence capable of expressing mature human fibroblast interferon.

B. "Ptashne Lab Methods"

1. Introduction

60. By March 19, 1980, a number of reports had been published that at least purported to describe generalized methods for the heterologous expression of genes in *E. coli*. In particular, the art available as of March 19, 1980, included three publications and one patent from the laboratory of Dr. Ptashne (Backman *et al.*, "Maximizing gene expression on a plasmid using recombination *in vitro*," Cell 13: 65-71 (1978) (Exhibit 2095), Roberts *et al.*, "A general method for maximizing the expression of a cloned gene," Proc. Nat'l Acad. Sci. USA 76: 760-764 (1979) (Exhibit 1126), Roberts *et al.*, "Synthesis of simian virus 40 t antigen in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 76: 5596-5600 (1979) (Exhibit 2093), and U.S. Patent No. 4,332,892 (Exhibit 2004)), the collective teaching of which will be referred to herein as the "Ptashne lab methods." Although the '892 patent did not issue until June 1, 1982, I have been advised that it should be included in this analysis because, as I have been informed, for certain purposes, the '892 patent is deemed to have an "effective date" of January 15, 1979. As demonstrated below, even the collective disclosure of these four reports from the Ptashne laboratory did not provide those of ordinary skill as of March 19, 1980 with a reasonable expectation that they could construct a DNA encoding the mature form of human fibroblast interferon having a total of 166 amino acids and unaccompanied by its signal peptide or that they

could express biologically active mature human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide.

61. The Ptashne lab methods relate to the assembly of a "hybrid ribosome binding site" that could be used for expression of prokaryotic, and perhaps eukaryotic, proteins in *E. coli*. These methods provide a conceptual approach that might be useful for heterologous expression of particular proteins in *E. coli* under fairly narrow, well defined circumstances. However, this approach possessed a number of limitations and a lack of guidance in particular aspects of the methods that precluded its general application. In fact, a later publication from the same laboratory (Guarente *et al.*, "Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin," Cell 20: 543-553 (1980) (Exhibit 2132)) recognized those limitations and provided a new approach that involved a highly effective screening method that permitted the detection of extremely rare clones capable of expressing a target protein. This method was used by Dr. Taniguchi for expression of the mature form of human fibroblast interferon (Exhibit 1009). Each of the four publications describing the Ptashne lab methods is discussed below.

2. Backman *et al.*, "Maximizing gene expression on a plasmid using recombination *in vitro*," Cell 13: 65-71 (1978) (Exhibit 2095)

62. In the introduction, Backman reviews the teaching of the art regarding the apparent requirement for a "Shine-Dalgarno" sequence (consisting of from 3 to 9 nucleotides complementary to the 3'-end of 16S RNA) disposed approximately 3-12 nucleotides upstream of the ATG (or GTG) initiation codon of the target protein to be expressed (Exhibit 2095, at 65, left col., first ¶). At page 70 (Exhibit 2095, at 70, left col., third full ¶), Backman discloses the construction of pKB280, which is depicted schematically in Fig. 2 (Exhibit 2095, at 68). The essential elements of this construction were that a 95-base-pair EcoRI-AluI restriction fragment

containing the *lac* promoter, operator, and SD sequence was directly joined to a restriction fragment carrying the ATG initiation codon for λ repressor preceded by 3 base pairs. The recombinant molecule formed included the *lac* SD sequence situated 8 base pairs upstream from the initiation codon of the λ repressor, and allowed expression of substantial levels of the target repressor protein (Exhibit 2095, at 67, Table 1). No "trimming" of either DNA molecule was employed in the preparation of the plasmids described in this publication (although a single base pair was lost for unknown reasons; Ex. 2095, at 95, right col., last ¶, first sentence). Backman speculated that if the ATG initiation codon were not adjacent to a restriction site, "it should be possible to make the proper length molecules starting with longer fragments and trimming with exonucleases and single-strand-specific nucleases, or by synthesis in vitro of DNA molecules containing an initiator ATG and adjacent sequences which may be placed between an appropriate SD sequence and the remainder of the gene" (Exhibit 2095, at 69, right col., first ¶).

3. U.S. Patent No. 4,332,892 to Ptashne *et al.* (Exhibit 2004)

63. The '892 patent is directed to a method for expressing proteins in their "native" form, *i.e.*, unfused to amino acid sequences of *another* protein. The method involves cloning a target gene near a restriction site, digesting away the DNA between that restriction site and the naturally occurring initiation codon of the coding sequence of the target gene, and then inserting a "portable promoter" upstream of that initiation codon (Exhibit 2004, at col. 1, ll. 35-48). The portable promoter includes a "Shine-Dalgarno" ("SD") sequence but does not carry an initiation codon (Exhibit 2004, at col. 1, ll. 35-48). The combination of the bacterial SD and ATG of the target protein is referred to as a "hybrid ribosome binding site." The '892 patent includes one "Example," which appears to be a description of how one might attempt to express rabbit β-globin in *E. coli* (Exhibit 2004, at col. 2, ll. 15-63). According to the Example, the rabbit β-globin gene is inserted into a plasmid vector in such a manner that there are unique EcoRI and

HindIII restriction sites approximately 55 and 25 base pairs, respectively, upstream of the naturally-occurring ATG initiation codon of β-globin. The '892 patent proposes that the plasmid could be linearized by HindIII digestion, and ExoIII and S1 nucleases used to remove DNA between the HindIII site and the ATG of the β-globin gene (Exhibit 2004, at col. 2, ll. 51-56). Since these enzymes would digest the DNA in both directions, the sequences between the HindIII and EcoRI sites would be removed as well. Therefore, molecules, in which the ATG was removed by nuclease treatment, would likely no longer contain the EcoRI site used to introduce the portable promoter. This arrangement of elements therefore effectively eliminates an entire population of non-expressing clones, thereby simplifying identification of desired clones. The Example further suggests that after digestion of the trimmed DNA with EcoRI, a 95-base pair EcoRI-AluI portable *lac* promoter fragment could be inserted to provide recombinant molecules potentially capable of expressing rabbit β-globin (Exhibit 2004, at col. 2, ll. 36-43). The only method described in the '892 patent for detection of heterologously expressed rabbit β-globin relies on the use of an RIA (Exhibit 2004, at col. 2, ll. 44-47). However, a subsequent report from the same laboratory (Guarente *et al.* (Exhibit 2132)) indicates that an RIA assay for rabbit β-globin in an *E. coli* extract apparently was impossible to do because "affinity-purified serum to rabbit β-globin cross-reacts in a solid phase radioimmune assay with at least one *E. coli* protein." Exhibit 2132, at 544, left col., first full ¶, lines 5-9. Moreover, the '892 patent does not provide experimental detail or, in fact, any information that the would indicate the experiments underlying Example had actually been carried out.

64. As of March 1980, it was known that some normally secreted proteins are encoded as a larger precursor protein, which includes an amino-terminal amino acid sequence referred to as the leader peptide. Processing of the precursor protein to remove the leader

peptide provides the mature, secreted form of the protein. Accordingly, the amino-terminal amino acid of the mature form is not encoded by an "initiation codon." The '892 patent does not disclose any information regarding expression of such a "mature" protein. This omission in the description of the '892 patent is not unexpected since the '892 patent explicitly provides that that portable promoter does not include an ATG initiation codon, and there is no *a priori* expectation that the amino-terminal amino acid of a mature protein would be methionine. In particular, in the present context, as of the January 15, 1979 effective date of the '892 patent, it is my understanding that there was no information in the art regarding the amino-terminal amino acid sequence of the mature form of human fibroblast interferon.

65. The expression of the mature form of human growth hormone illustrates the limitation to the "Ptashne lab method" of U.S. Patent No. 4,332,892 (Exhibit 2004), discussed in the preceding paragraph. The method described in '892 patent refers to expression mediated by a "hybrid ribosome binding site" in which the Shine-Dalgarno sequence ("SD sequence") is derived from the *lac* operon of *E. coli* while the ATG initiation codon is that of the native form of the target protein, *i.e.*, unfused to any other protein. That is, the method of the '892 patent, as well as, in fact, that of any of the other publications describing the "Ptashne lab methods," makes no provision for direct expression of a coding sequence, such as that for human growth hormone, that begins with phenylalanine.

4. Roberts *et al.*, "A general method for maximizing the expression of a cloned gene," Proc. Nat'l Acad. Sci. USA 76: 760-64 (1979) (Exhibit 1126)

66. Notwithstanding this publication's title, the data provided clearly demonstrate only that expression of a cloned gene is highly dependent on the specific nature of the upstream, non-translated mRNA. It is also apparent that the variation in expression of the target protein could not be directly correlated to the "positioning" of the *lac* SD sequence in relation to the *cro*

SD sequence and ATG initiation codon of that target *cro* protein. In this, context, reference to "*lac-cro* fusions" is somewhat of a misnomer. Although clones with "enormous" differences in expression levels were reported in this publication, the authors could not provide any predictable approach to the expression of heterologous proteins in *E. coli*, *i.e.*, "[w]e have no explanation for these differences" (Exhibit 1126, at 763, right col., last ¶, last full sentence).

67. More specifically, Roberts disclosed the cloning of the gene encoding the *cro* protein of the *E. coli* bacteriophage λ in such a manner that a unique BamHI restriction site was disposed some 54 base pairs upstream of the ATG initiation codon (pTR151). It is critically important to note that the *cro* gene in that restriction fragment carried its own SD sequence and ATG initiation codon; in fact that the *cro* SD is a relatively long, nine-base pair (TAAGGAGGT) match for the RNA sequence at the 3'-end of 16S rRNA (Exhibit 1126, at 763, right col., first full ¶). The resulting plasmid (pTR151) was linearized by digestion with BamHI, treated sequentially with Exonuclease III and S1 to remove DNA between the BamHI site and the ATG initiation codon of *cro*, and a 95-base pair fragment inserted, which fragment carried the *lac* promoter and the *lac* SD sequence (AGGA). As discussed below, these procedures were expected to produce a large population of clones, only a very small proportion of which would be of interest. However, in this instance, three layers of screening methods allowed the rapid and facile identification of the nine, different specific clones reported. The first screening layer was a colorimetric plate screen that identified those clones carrying the inserted *lac* promoter. This screen is based on the observation that the *lac* promoter fragment also carries the *lac* operator sequence, which is bound by *lac* repressor. Transformants carrying the *lac* promoter-operator effectively titrate the limited number of repressor molecules in the cell and thereby induce expression of the gene for β-galactosidase. By plating the putative transformants on selective

media including the β-galactosidase chromogenic substrate 5-bromo-4-chloro-3-indolyl-β-D-galactoside ("XG"), clones carrying the *lac* promoter-operator therefore are readily identified by the production of the deep blue pigment, indigo. The second layer of screening was based on the biochemical properties of the target *cro* protein itself, *i.e.*, clones expressing *cro* are immune to infection by the bacteriophage λ. Synthesis of the target protein (*cro*) was "assayed" by "cross-streaking" putative transformants against a suspension of bacteriophage λ. In fact, the authors predicted (accurately) that as the level of *cro* production increased, the host *E. coli* would become immune to infection by ever more virulent derivatives of bacteriophage λ. Using these two plate screens, the pool of transformants obtained was readily divided into a small number of sub-groups. Plasmid DNA was isolated from 40 clones representing all of the identified sub-groups, digested with the restriction endonuclease HaeIII, and analyzed by acrylamide gel electrophoresis. Using information from this third level of screening, nine distinct, representative transformants were chosen for the detailed analyses reported (Exhibit 1126, at 762, right col., first two ¶¶ in the "Results" section).

68. In five of the clones analyzed (*see* Exhibit 1126, at 763, Fig. 3) neither the *cro* SD sequence, nor its relation to the ATG initiation codon were altered in any way, even though the level of expression of *cro* protein varies by a factor of approximately 20 between clones. This is noteworthy because, as acknowledged by the authors, translation initiation is expected to be "driven" by recognition of the *cro* SD sequence and the naturally-occurring ATG initiation codon of *cro*. Accordingly, it appears that the relative positioning, or even the presence of the *lac* SD sequence, has no bearing on the level of expression of *cro* protein.

69. In a sixth clone (pTR182), the manipulations discussed above generated a DNA molecule in which a single nucleotide was deleted from the *cro* SD sequence. This clone

exhibited a level of expression of *cro* that is approximately one half the maximum observed. However, deletion of almost all (pTR182) or all (pTR190) of the *cro* SD sequence, while placing the *lac* SD sequence 10 base pairs and 5 base pairs, respectively, upstream of the *cro* effectively *eliminated* expression (Exhibit 1126, at 763, right col., last 2 lines to 764, left col., first 3 lines).

70. As noted above, in five of the recombinant plasmids constructed, *i.e.*, those designated pTR213, pTR199, pTR214, pTR188, and pTR194, neither the *cro* SD nor its relation to the *cro* ATG initiation codon have been altered. Therefore, in each instance transcription of the *cro* gene is driven by the same *lac* promoter and translation initiation of the mRNA produced is driven by the same SD sequence and ATG initiation codon. In light of these observations, these five clones actually represent different deletion mutations in the 92-base pair non-coding sequence of the mRNA, upstream of the SD and ATG regulatory signals. It is therefore apparent, as acknowledged by the authors, that the enormous variations in observed expression of *cro* protein are due to "some post-transcriptional process." Exhibit 1126, at 763, right col., ll. 16-20. Although the authors speculated that such processes could be related to mRNA stability, mRNA processing, or ribosome binding efficiency, they acknowledged that they did "not know how the presence of different leaders might influence the stability or processing of a message." Exhibit 1126, at 763, right col., ll. 23-24.

71. It has been suggested in the art that the data of Roberts may reflect the influence of mRNA stability in *E. coli* as well as the effects of the secondary and tertiary structure of the mRNA (determined by the mRNA nucleotide sequence) on protein expression in that host organism (*see, e.g.*, Steitz (Exhibit 1109), at 492, first ¶)).

72. It is my opinion that the experiments described and the data reported in Roberts demonstrated the variable, unpredictable, and protein-specific nature of heterologous gene

expression in *E. coli*. That is, the experiments described and the data reported in Roberts were unrelated to "positioning" of the *lac* ribosome binding site or the generation of "*lac-cro*" fusions. This is particularly apparent since the two isolates using the *lac* SD for *cro* expression, pTR182 and pTR190, produce, respectively, barely detectable and undetectable levels of *cro* protein (Exhibit 1126, at 763, right col., last 2 lines to 764, left col., first 3 lines). Moreover, it is apparent that completely different expression data would be obtained using the same upstream deletion mutants disclosed by Roberts, but joined to a coding sequence for a protein other than *cro* (Steitz (Exhibit 1109), at 492, first ¶). Accordingly, Roberts teaches that expression of an *E. coli* gene, which already carries an appropriately positioned ribosome binding site (SD and ATG sequences) will be unpredictably sensitive to the distance separating the promoter and the gene, *i.e.*, expression will be sensitive to the length, nature, and structure of the non-coding mRNA upstream from the SD sequence and the ATG initiation codon.

5. Roberts *et al.*, "Synthesis of simian virus 40 t antigen in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 76: 5596-5600 (1979) (Exhibit 2093)

73. Roberts discloses expression of SV40 t antigen, which is a cytoplasmic viral protein, in *E. coli*. In this work, Roberts inserted a gene encoding the target protein into the *E. coli* plasmid vector, pBR322, such that the ATG initiation codon for the SV40 t antigen was only 7 base pairs from a unique HindIII site. The plasmid was opened at the HindIII site, digested sequentially with Exo III and then S1 to remove some or all of this 7 base pair sequence. The "trimmed" DNA was then digested with EcoRI, a DNA fragment carrying the *lac* promoter and SD sequence was inserted, and the pool of recombinant plasmids was used to transform an *E. coli* host strain. Roberts reported that approximately half of the clones obtained produced a 20 kD protein that was precipitated by anti-t antigen antisera. In particular, a clone in which a total of 4 of the 7 base pairs had been removed (pTR436) exhibited the highest levels of

synthesis of a 20 kD protein presumed to be the SV 40 t antigen. The starting material (pTR422), *i.e.*, without any Exo III or S1 digestion, provided a detectable level of target protein. Another clone (pTR440), in which a total of 2 of the 7 base pairs had been removed, and therefore differed only by 2 base pairs from pTR436, produced much less of the 20 kD protein than was observed with pTR436 (Exhibit 2093, at 5598, Fig. 4).

74. The data provided (Fig. 4) also demonstrate that apparently there is a methionine codon (*i.e.*, internal to the SV 40 t antigen coding sequence) that is efficiently recognized as an initiation codon in *E. coli* and directs the synthesis of an approximately 10 kD molecular weight polypeptide (Exhibit 2093, at 5598, Fig. 4). In fact, the stained bands of Fig. 4 suggest that at least as many, if not more, copies of the 10kD protein are synthesized as compared to the number of presumed full length SV40 t-antigen, 20 kD polypeptides produced by pTR436 (Exhibit 2093, at 5599, right col., first ¶, last sentence).

75. It is my opinion that the experiments disclosed in this reference provide a very narrowly-circumscribed, favorable demonstration of the feasibility of placing a "portable promoter" adjacent to the ATG initiation codon of a mammalian coding sequence to facilitate heterologous expression of the encoded protein in *E. coli.* In this instance, as noted above, a unique restriction site (AAGCTT) occurs naturally only 7 base pairs upstream of the ATG initiation codon for the target protein (**AAGCTT**TGCAAAG**ATG**). A careful review of Fig. 4 at page 2 indicates that only 2 nucleotides needed to be excised by the Exo III treatment, leaving a 6-nucleotide, single-stranded end to be removed by the subsequent S1 digestion, to provide the clone designated pTR440. Similarly, only 4 nucleotides were required to be excised by Exo III treatment, leaving an 8-nucleotide, single-stranded end to be removed by the subsequent S1 digestion, to provide the clone designated pTR436. In this instance, *i.e.*, where so few base pairs

were to be removed, "brute force" screening methods could be applied to detect SV40 small t antigen expressing clones since only a very limited digestion of the DNA was required (*i.e.*, removal of more than 8 nucleotides by Exo III would destroy the t antigen initiation codon). Under such favorable circumstances, it was found that "approximately half the clones produced a protein of approximately 20 kd, which was specifically precipitated by anti-T sera" (Exhibit 2093, at 5598, right col., first ¶, ll. 2-4). However, as understood by those of ordinary skill in the art, these results could not be extrapolated to an instance in which, for example, 70 base pairs had to be removed in order to reach the ATG initiation codon of the mammalian gene to be expressed in *E. coli*. In this context, note the 500-fold difference in the frequency reported for expression of the precursor form of human fibroblast interferon (where the starting point for nucleolytic digestion was 10 base pairs from the ATG of the precursor protein) as compared to that for the mature form of human fibroblast interferon (where the starting point for nucleolytic digestion was 70 base pairs from the ATG of the precursor protein) (*i.e.*, *see* Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009), at 5232, left col., first full ¶). The practical consequence of this immense difference in frequency is that, in my opinion, it becomes unduly burdensome or impossible to rely upon "brute force" methods to identify the desired clones (*see* Guarente *et al.*, "Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin," Cell 20: 543-553) (Exhibit 2132).

6. U.S. Patent No. 4,342,832 to Goeddel *et al.* (Exhibit 1058)

76. In the '832 patent, Goeddel discloses, *inter alia*, reagents and methods used for high-level expression of the mature form of human growth hormone ("hGH"). Goeddel suggests, in general terms and without providing any experimental detail, that a signal peptide might be removed by the combined activity of either "an exonuclease" or a DNA polymerase (in

the presence of a deoxyribonucleoside triphosphate) followed by treatment with S1 nuclease to remove single stranded "tails" generated by the exonuclease or polymerase (Exhibit 1058, at col. 5, l. 60 through col. 6, l. 27). It should also be noted that the hypothetical sequence to be removed in that discussion involved only ten base pairs. However, even in that instance, one skilled in the art would infer that the proposed use of an exonuclease followed by S1 treatment would not provide a precise, well controlled approach for the purpose intended. This is so because, after mentioning the use of a nuclease followed by S1, the '832 patent states: "Alternatively, *and more precisely*, one may employ DNA polymerase digestion in the presence of deoxynucleotide triphosphates," generally according to the method described and the reference cited (Exhibit 1058, at col. 6, ll. 15-27). In fact, Goeddel did not use either approach in the disclosed and claimed invention. The only use noted for nuclease S1 was for removal of the four nucleotide long single stranded ends produced by digestion by the restriction endonuclease EcoRI (Exhibit 1058, at col. 11, ll. 53-58). Therefore, it is my opinion that those of ordinary skill in the art as of March 19, 1980 would not have concluded that '832 patent provided sufficient guidance, *e.g.*, for removal of a 70-base pair leader sequence using a combination of "an exonuclease" and S1.

IV. PROBLEMS INHERENT TO TAILORING METHODS AS OF MARCH 19, 1980

77. As of March 19, 1980, those of ordinary skill in the art recognized that heterologous expression of a mammalian protein in *E. coli* required that the DNA encoding the target protein be "tailored" to remove DNA sequences upstream of the ATG initiation codon to allow insertion of a SD sequence (Steitz (Exhibit 1109), at 490, last ¶, which ends on page 491). As of that date, those of ordinary skill in the art also recognized although the "Ptashne lab methods" might be effective in instances in which only a short double-stranded DNA segment was to be removed in the "tailoring" process, but those methods would not be useful where, *e.g.*,

70 base pairs had to be removed. This was so because the methods inherently yield a large population of DNA molecules, only a fraction of which represent the desired construction. Moreover, both the size of the population and the range of molecular products appeared to increase, seemingly exponentially, as the length of the segment to be removed increased. Accordingly, the frequency with which desired clones could be found fell dramatically as the length of the segment to be removed increased (*e.g.*, a five-hundred fold decrease in the frequency with which desired clones were identified as the size of the segment to be removed inceased from 10 to 70 base pairs, *i.e.*, a factor of seven (Taniguchi *et al.* (Exhibit 1009), at 5232, left col., first full ¶). These observations are not surprising in view of the number and the nature of the steps involved in the Ptashne lab methods, as well as the properties of the nucleases used.

78. As described above regarding the '892 patent and Roberts *et al.* (Exhibit 2093), the Ptashne lab methods for tailoring of a previously-cloned DNA sequence to remove unwanted DNA segments and to introduce a "portable promoter" involved a total of eight steps: (1) digestion of the DNA with a first restriction enzyme to provide a starting point, (2) removing one strand of the duplex DNA on both sides of the starting point by digestion with nuclease Exo III, (3) removal of the resulting single-stranded "tails" with nuclease S1, (4) digestion of the resulting population of molecules with a second restriction enzyme, (5) ligating a "portable promoter" DNA fragment (with appropriate ends) to the population of molecules from the prior step, (6) transforming an appropriate *E. coli* host cell with that population of recombinant molecules, (7) selecting the transformants carrying a recombinant plasmid from within the population (generally by selection for the antibiotic resistance marker carried by the expression vector), and (8) identifying clones carrying the desired constructs in which the target protein is expressed. In some instances the second and third steps were, in essence, combined by use of

nuclease Bal 31 (rather than sequential treatment with Exo III and S1) which can remove both strands from the end of a duplex DNA molecule. However, by analogy to multi-step chemical syntheses, if the yield at each step of the process is, *e.g.*, 25%, the overall yield of an eight step process is $(.25)^8 = 0.0015\%$. Therefore, those of ordinary skill in the art recognized that all the steps of a multi-step process would have to work well for the overall process to provide the desired clones at a detectable frequency. Thus, success with one or more individual steps, in isolation, would not convince those skilled in the art that the "Ptashne lab methods" could be generally applied with a reasonable expectation of success. This conclusion was particularly apparent in an "extreme" case, *e.g.*, the proper tailoring of the gene encoding the precursor form of human fibroblast interferon to provide a DNA encoding the mature form of human fibroblast interferon having a total of 166 amino acids and unaccompanied by the coding sequence for its signal peptide.

79. The Ptashne lab methods include the use of six different enzymes: the first and second restriction enzymes, T4 DNA ligase, and, of particular concern, the nucleases Exo III, S1, and Bal 31. These nucleases were used for the very precise digestions of one or both strands of DNA molecules that are inherent to the Ptashne lab methods. However, these enzymes are not ideally suited for these purposes, *e.g.*, the "single-strand specific" nuclease S1 is capable of digesting both strands of a duplex DNA molecule. Moreover, the rate of digestion by Exo III Bal 31, and S1 was expected to depend upon the source and purity of the enzyme, the nature of the substrate DNA, and the specific reaction conditions employed. A subsequent report (Wei (Exhibit 1123)) demonstrated that preparations of Bal 31 included variable proportions of two forms of the nuclease, each of which had different kinetic properties (Exhibit 1123, at 13509, Table III and 13510, Table IV). In addition to the above, any untoward damage to the substrate

DNA, both as a result of the inability to control the activity of each of these enzymes tightly and reproducibly and as a result of the activity of contaminating enzymes, would increase as a function of the amount of time the DNA was exposed to each enzyme. This could, for example, lead to a bias in favor of readily-ligated, undigested and undamaged starting material over non-ligatable, damaged DNA fragments. In this context, note that even under the "best case," circumstances reported from the Ptashne lab (*see* Roberts *et al.* (Exhibit 2093)), although half of the clones produced the target protein, half of the isolated clones did not produce immunoprecipitable SV40 small t antigen. This is noteworthy since the undigested starting material (pTR422, which was not exposed to either ExoIII or S1) was reported to direct expression of a "detectable" amount of SV40 small t antigen (Exhibit 2093, at 5598, right col., l. 9 to 5599, right col., l. 24). Therefore, in view of all of these elements, as of March 19, 1980, it is my opinion that those of ordinary skill did not expect that the Ptashne lab methods could be directed successfully to the tailoring of a DNA molecule encoding the precursor form of human fibroblast interferon, since the enzymes used could not be adequately controlled for the time required to remove nucleotides encoding the human fibroblast interferon signal peptide.

80. As of March 1980, those of ordinary skill understood that DNA digestions using the nucleases of the Ptashne lab methods were not easily controlled, particularly where a longer stretch of DNA was to be removed and a precise end point desired. For example, nuclease S1 is capable of degrading double-stranded DNA (*e.g.*, *see* Bendig *et al.*, "Deletion mutants of polyoma virus defining a nonessential region between the origin of replication and the initial codon for early proteins," J. Virol. 32: 530-535 (1979) (Exhibit 1127), at 553, Fig. 3 and 532, left col. first ¶, last sentence) demonstrating that S1 alone can remove up to 89 base pairs of a duplex DNA molecule, and Shenk *et al.*, "Biochemical method for mapping mutational alterations in

DNA with S1 nuclease: The location of deletions and temperature-sensitive mutatiosn in simian virus 40," Proc. Nat'l Acad. Sci. USA 72: 989-993 (1975) (Exhibit 1128), which reported that the "nibbling" activity of S1 nuclease was capable of removing about 30 base pairs from the end of a double-stranded DNA molecule (Exhibit 1128, at 990, right col. first ¶) and that this "nibbling" effect was greater with a commercial source of S1 as compared S1 purified as described (Exhibit 1128, at 989, right col., third ¶)).

81. As another example of the ability of S1 to "trim" double-stranded DNA, it was observed in the Fiers laboratory (*see* Remaut *et al.*, "Inducible high level synthesis of mature human fibroblast interferon in *Escherichia coli*," Nucl. Acids Res. 11: 4677-4688 (1983) (Exhibit 1029)) that S1 "nibbling" apparently was responsible for the removal of 3 base pairs of duplex DNA (Exhibit 1029, at 4682, second ¶, last sentence)).

82. The Ptashne lab methods, as described, were directed toward the use of the *lac* portable promoter, an obvious choice since that DNA fragment was available and carried a SD sequence only five base pairs from the end of the molecule. However, those of ordinary skill would have been concerned that constitutive expression of the very hydrophobic, nonglycosylated mature form of human fibroblast interferon would be toxic to the host cell. Accordingly, this toxicity would provide a negative selection against the very clones sought, *i.e.*, those (constitutively) expressing mature human fibroblast interferon. In fact, this initial concern was subsequently confirmed (Remaut *et al.* (Exhibit 1029), at 4682 and 4687).

83. A fundamental limitation to the Ptashne lab methods was the absence of a detection method that would allow the rare, desired, and appropriately tailored clones to be identified within the large population generated. For the reasons noted above, the use of the Ptashne lab methods for tailoring of a DNA encoding the precursor form of human fibroblast

interferon was expected to generate a very large population of clones in which the desired isolate would represent only a very small fraction. That is, as of March 19, 1980, for the reasons provided above, extended digestions according to the Ptashne lab methods were expected to provide a distribution of recombinant molecules that could be depicted or envisioned as a very flattened bell curve. Notwithstanding this observation, where the target protein had an activity that can be readily identified, the size of that population generated by the Ptashne lab methods is not necessarily fatal. For example, application of three sequential screening methods allowed the stepwise narrowing of the population of isolates and the ultimate identification of clones expressing various levels of the prokaryotic *cro* protein (which provides resistance to bacteriophage λ) with minimal effort. In contrast, in the absence of viable screening methods, it is my opinion that identification of rare isolates expressing mature human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide within an immense population of clones would have been unduly burdensome or impossible as of March 19, 1980. This conclusion is supported by the actual frequency ultimately reported, in October 1980, from the Ptashne lab, *i.e.*, desired clones were found at a frequency of 0.01% (Taniguchi *et al.* (Exhibit 1009), at 5232, left col., first full ¶, last sentence).

84. It is my opinion that, as of March 19, 1980, those of ordinary skill would not have reasonably expected that clones producing mature human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, could be identified within the population of isolates expected to be generated by the Ptashne lab methods using "brute force" analyses, without undue experimentation. Examples of such analyses would include, for example, isolation, characterization, and ultimately, sequencing of plasmid DNA from random clones. Although optimized methods, reagents, and even automated instruments for these

processes are available today, in March of 1980 each of these steps was labor-intensive (*e.g.*, *see* Chapter 5: "Antigen-Antibody Interactions," PRINCIPLES OF IMMUNOLOGY, 2nd Ed., MacMillan Publishing Co., New York (1979) (Exhibit 1016), at 65-79). Such "shotgun" approaches could reasonably be used where the expected frequency of desired clones would be very high. However, it is my opinion that those of ordinary skill in the art would not have considered, much less attempted, reliance on such brute force, shotgun approaches for the analysis of thousands of potential isolates. Moreover, even if there were any motivation to embark on that endeavor, it would have been severely mitigated in view of the likelihood that expression of hydrophobic, nonglycosylated mature human fibroblast interferon could be toxic to the host cell. Moreover, as I recall, a fibroblast interferon-specific RIA had not been established as of March of 1980. Similarly, the preparation and analysis of extracts for assay of anti-viral activity (CPE) from thousands of clones would entail an undue amount of experimentation as of March 19, 1980.

85. Therefore, it is my opinion that as of March 19, 1980, those of ordinary skill in the art understood that even if there were a basis for using Ptashne lab methods to tailor the DNA encoding the precursor form of human fibroblast interferon, it would not have been possible to identify the desired clones expressing the mature form of that protein. In fact analysis of each of the reports published after March 19, 1980, demonstrating expression of mature human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide, reveals that the desired recombinant molecules (a) were constructed using methods expected to yield a single recombinant molecule (Goeddel (Exhibit 1012)), (b) were identified within an immense population using an effective screening method (Taniguchi *et al.* (Exhibit 1009)), or (c) were obtained using a combination of both approaches (Remaut *et al.* (Exhibit 1029)).

V. IMPEDIMENTS TO HETEROLOGOUS GENE EXPRESSION IN *E. COLI* IN ADDITION TO GENE TAILORING METHODS

86. It was also recognized by March 1980, that *even if* a mammalian gene could have been properly tailored for expression, it was, nevertheless quite possible that the target protein would not be produced in a biologically active form in *E. coli*. This was particularly apparent in the present instance, in which the target protein is a normally secreted mammalian glycoprotein, *i.e.*, the nonglycosylated mature form of human fibroblast interferon having a total of 165 or 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide. Expression of biologically active mature human fibroblast interferon could be obviated by instability of the encoding mRNA, impaired translation of that mRNA resulting from, *e.g.*, inhibitory secondary structure formation or biased codon usage in the mammalian coding sequence (compared to that for *E. coli*), instability of protein, inability of the non-glycosylated mature form of human fibroblast interferon unaccompanied by its signal peptide to fold properly within an *E. coli* host cell, and finally, the potentially toxic effects of expression of this very hydrophobic protein in *E. coli*.

87. As of March 19, 1980, researchers in the art of molecular biology were seriously concerned that attempts to produce any *particular* mature mammalian protein (and particularly any particular normally secreted mammalian glycoprotein encoded with a signal peptide) in bacteria would be fraught with problems. This is because bacteria simply do not ordinarily produce mammalian or other eukaryotic proteins and, although there are common features to bacterial and mammalian physiology and genetics, major differences were assumed to confer an unpredictability to efforts to express specific mammalian proteins in bacteria. Thus, for example, one successful direct expression of a mature mammalian protein in bacteria, namely, mature human growth hormone by David Goeddel in 1979, did not provide a level of comfort

and reasonable expectation of success to one of ordinary skill that the approach of the Goeddel hGH paper, or a variant thereof, or any other method would predictably result in the production of any other specific mammalian protein, such as non-glycosylated human fibroblast interferon, in biologically active form.

88. Briefly, one skilled in the art of molecular biology and biochemistry during 1979 and through March 19, 1980 would have expected to encounter difficulties in expressing any specific mammalian protein in *E. coli* (and particularly any particular normally-secreted mammalian glycoprotein encoded with a signal peptide) at one or several levels of the steps that are naturally required to express the protein of interest. These include:

(a) inability of the bacterial cell to transcribe the mammalian gene;

(b) instability of mRNA encoding the mammalian protein in the bacterial cell;

(c) impaired translation of the mRNA, particularly due to differences in codon selection between bacteria and mammalian systems as well as the sequence-dependent mRNA secondary and tertiary structure;

(d) instability of the mammalian protein in the bacterial environment;

(e) inability of the bacterial cell to fold the mammalian protein into its proper three-dimensional conformation, particularly in those instances in which the protein is normally secreted and forms structure-stabilizing disulfide bonds;

(f) lack of or inappropriate glycosylation of the protein, as required for protein stability and biological activity; and

(g) potential toxic effects of the mammalian protein on the bacterial metabolism, thus preventing production of the protein in any meaningful quantities as well as making it markedly more difficult to construct and identify bacteria expressing

that mammalian protein, since the desired isolate would be at a selective disadvantage.

89. Most concerns related to steps (c)-(g) and were based on considerable scientific evidence.

90. For any given mammalian protein, any one impediment of the above, or several together, could have led to failure to produce a biologically active protein in bacteria. These potential impediments to expression of a given mammalian protein in bacteria would exist for any expression method, including the method of the Goeddel hGH paper (Exhibit 1057) or any other method.

91. During 1979 and through March 19, 1980, researchers were concerned that any particular mammalian mRNA would be unstable in the environment of the bacterial host and this would prevent its expression. Indeed, significant functional, structural and chemical differences exist between mammalian and bacterial mRNAs in their native environments, and efficient degradation enzymes (nucleases) that rapidly degrade mRNA exist in bacteria. For example, Hautala *et al.* concluded in "Increased expression of a eukaryotic gene in *Escherichia coli* through stabilization of its messenger RNA," Proc. Nat'l Acad. Sci. USA 76: 5574-5578 (1979) (Exhibit 1110) that there are inherent structural differences between bacterial and mammalian mRNAs. They suggested that "eukaryotic mRNAs may possess some unique secondary or tertiary structure not present in prokaryotic mRNAs . . ." Exhibit 1110, at 5578.

92. During 1979 and through March 19, 1980, it was known that mammalian mRNAs differed considerably from bacterial mRNAs in their codon usage:

(a) Certain codons are favored in *E. coli* mRNAs for some amino acids while different codons are favored in mammalian mRNAs, and these differences are

reflected in the composition of the pools of tRNAs in these different organisms. See, for example, Itakura *et al.*, "Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin," Science 198: 1056-1063 (1977) (Exhibit 2080); Fiers *et al.*, "Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene" Nature 260: 500-507 (1976) (Exhibit 1115); Grantham *et al.*, "Codon catalog usage and the genome hypothesis," Nucl. Acids Res. 8: r49-r62 (1980) (Exhibit 1129); and Fraser and Bruce, "Chicken ovalbumin is synthesized and secreted by *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 75: 5936-5940 (1978) (Exhibit 2084, at 5940, left col., ll. 1-4);

(b) Additionally, as noted above, bacterial mRNAs naturally include a sequence, known as the "Shine-Dalgarno sequence," in the 5' untranslated leader region of the mRNA, which is required for efficient bacterial ribosome binding and initiation of translation. Shine and Dalgarno, "Determinant of cistron specificity in bacterial ribosomes," Nature 254: 34-38 (1975) (Exhibit 1043). Thus, mammalian mRNAs had to be engineered to encode a Shine-Dalgarno sequence in a proper sequence context; and

(c) Considering the substantial differences in codon usage, it was thought that, even if an appropriate "Shine-Dalgarno" sequence context were to be introduced, this would not guarantee efficient translation of the bacterially encoded, mammalian mRNA. In addition to mediocre translation efficiency, the different codon usage and poor translation may also result in rapid degradation of foreign mRNA, which could vary from mRNA to mRNA. It was understood by persons of ordinary skill

as of March 19, 1980 that the differences in codon usage and the inability to provide a proper "Shine-Dalgarno" sequence context were potentially fatal to the ability of bacteria to translate that mRNA to produce the mammalian protein.

93. Even if there were no problems of mRNA instability and deficient translation, there was considerable concern that the desired mammalian protein may not be stable in the bacteria. See, for example, the report of expression of a small fusion protein containing the first seven amino acids of β-galactosidase fused to -Glu-Phe-Met- and immediately followed by the 14 amino acids of somatostatin in the paper by Itakura *et al.*, "Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin," Science 198: 1056-1063 (1977) (Exhibit 2080), at 1060:

> The DNA sequence analysis of pSOM1 indicated that the clone carrying this plasmid should produce a peptide containing somatostatin. However, to date all attempts to detect somatostatin radioimmune activity from extracts of cell pellets or culture supernatants have been unsuccessful. . . . In a reconstruction experiment we have observed that exogenous somatostatin is degraded very rapidly by *E. coli* RR1 extracts. The failure to find somatostatin activity might be accounted for by *intracellular degradation by endogenous proteolytic enzymes* (emphasis added).

While we know today that this rapid degradation may have been in part due to the small size of the somatostatin fusion protein, as of March 19, 1980, one skilled in the art would have been very concerned that any particular mammalian protein, including mature human fibroblast interferon, would be subject to degradation. As of March 19, 1980, the minimum requirements for stability of a foreign protein in a bacterial environment were not known.

94. Although proteolytic degradation in *E. coli* was not a well studied phenomenon as of March 19, 1980, it was well known then and it is still recognized that *E. coli* possesses an efficient protein degradation system which selectively degrades aberrant proteins and protein fragments. Goldberg, "Degradation of abnormal proteins in *Escherichia coli*," Proc. Nat'l Acad.

Sci. USA 69: 422-426 (1972) (Exhibit 1111). It has been shown that mammalian proteins, which have not evolved to exist in bacteria, are frequently labile in *E. coli*. Simon *et al.*, "Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 80: 2059-2062 (1983) (Exhibit 1112); Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009).

95. I am also aware of experiments published in September 1980 in which two closely related polypeptides, prefibroblast interferon and mature fibroblast interferon, were expressed in separate *E. coli* cultures. Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009). As shown by pulse-chase labeling experiments (Exhibit 1009, at 5232, Fig. 4), a protein indicated to be prefibroblast interferon was completely degraded within 50 minutes; a protein indicated to be mature fibroblast interferon was about 50% degraded within 50 minutes. Thus, even knowing that one form of a protein, *e.g.*, pre-protein, can be expressed and isolated does not allow one to predict the stability of another closely related protein, *e.g.*, the corresponding mature protein

96. Thus, for these reasons given in paragraphs 86 to 95 above, it is my opinion that one of ordinary skill in the art as of March 19, 1980 did not expect or predict that mature human fibroblast interferon could be produced in bacteria or that it would be stable in *E. coli*. For the reason under paragraph 92 alone one ordinarily skilled in the art would not have been able to reasonably predict, *e.g.*, that the approach of the Goeddel hGH paper or that of the "Ptashne lab methods," or variants thereof, would be useful to produce human fibroblast interferon in mature form unaccompanied by the corresponding signal peptide. In addition, even though I was in the

end successful at expressing interferon activity in *E. coli*, I was myself highly doubtful in 1979-1980 that fibroblast interferon could be expressed in a stable, biologically active form in *E. coli*.

97. Even if one were able to obtain expression of a particular mammalian protein in a stable form in *E. coli*, one could not predict whether that protein would be properly folded to a suitable three-dimensional configuration such that the final molecule exhibits biological activity. In addition, many eukaryotic proteins, including human fibroblast interferon, are N-glycosylated and bacteria do not have the ability to effect N-glycosylation. Considering actual and often assumed requirements of N-glycosylation for proper folding, stability and biological activity, it was a particularly great concern that naturally glycosylated proteins could not be expressed in bacteria in a properly folded, biologically active and/or stable form. This statement would be true no matter what approach or method was used to express the mammalian protein in bacteria.

98. The biochemical environment inside bacteria differs significantly from that of mammalian cells with respect to protein modification. Some post-translational modifications, *e.g.*, N-glycosylation, hydroxylation, carboxylation, tyrosine-phosphorylation, lipidation and methylation of proteins found in mammalian cells just do not occur at all or do not occur with similar specificity in the bacterial cell. Such post-translational modifications are often vital to biological activity of many mammalian proteins, and it was assumed that N-glycosylation was of great importance for the integrity and activity of fibroblast interferon. These mammalian post-translational events, which have been reviewed by Uren, "The recovery of genetically engineered mammalian proteins," Am. Biotechnol. Lab. 2: 51-54 (1983) (Exhibit 1113), should have been known to one of ordinary skill as of March 19, 1980.

99. The proteins produced by the direct expression method described in the Goeddel hGH paper (Exhibit 1057) are located in the cytoplasm within the inner membrane of the *E. coli*

cell. As of March 19, 1980, one skilled in the art should have known that the bacterial cytoplasm was a very different environment and whether and how it could substitute for the environment in which secreted proteins, such as fibroblast interferon, are released from mammalian cells was unclear at best.

100. One skilled in the art as of March 19, 1980 was also aware of the importance of proper protein folding to biological activity. See, for example, Epstein *et al.*, "The genetic control of tertiary protein structure: Studies with model systems," Cold Spring Harbor Symp. Quant. Biol. 28: 439-449 (1963) (Exhibit 1114). Because protein folding depends on the local environments, and because bacteria do not duplicate the natural folding environment found in mammalian cells, one skilled in the art as of March 19, 1980 would have considered this to present a serious impediment to obtaining biological activity for a given mammalian protein expressed in bacteria. Furthermore, since improperly folded proteins are highly sensitive to proteolytic degradation in bacteria, Simon *et al.*, "Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 80: 2059-2062 (1983) (Exhibit 1112), one skilled in the art would also have regarded this as another reason to expect stability problems for mammalian proteins expressed in bacteria.

101. Finally, the expressed mammalian protein could have toxic effects on the bacterial cell, which then would prevent production of the protein in any meaningful quantity by these bacteria. These concerns regarding heterologous expression of potentially-toxic proteins in *E. coli* were important in view of the use of the *lac* promoters for expression. For example, use of a the *lac* portable promoters from the Ptashne lab provided a number of advantages, including the opportunity to employ colorimetric plate screens to identify clones carrying plasmids into which the portable promoter fragment had been inserted. However, transcription from such

plasmid-borne *lac* promoters would be constitutive in a wild-type *E. coli* host cell and therefore would have been expected by those of ordinary skill as of March 19, 1980 to impose a negative selection against clones expressing a toxic protein. In retrospect, this was indeed the case for human fibroblast interferon expression, which, at least in some strains of *E. coli* strongly inhibited the growth of the bacteria, thus effectively preventing interferon expression. Indeed, Remaut *et al.*, "Inducible high level synthesis of mature human fibroblast interferon in *Escherichia coli*," Nucl. Acids Res. 11: 4677-4688 (1983) (Exhibit 1029) state that "even marginal expression of β-interferon led to a drastic reduction of the growth rate of the bacteria." Exhibit 1029, at 4682.

102. Thus, I conclude that, for the reasons given in paragraph 86-101, one skilled in the art as of March 19, 1980 did not expect an N-glycosylated, mammalian protein to be expressed in a properly folded configuration with biological activity in bacteria. Around that time, I, for sure, was highly doubtful that even if I would be able to express human fibroblast interferon in *E. coli*, that it would be produced in a form that would be biologically active, similarly to naturally glycosylated, mature human fibroblast interferon.

VI. THE SCREENING METHOD OF GUARENTE (AFTER MARCH 19, 1980)

103. In June of 1980, Guarente (the Ptashne laboratory) published a report describing a newly-developed, highly-efficient screening method. This method was designed to facilitate identification of recombinant molecules carrying the amino-terminal coding region of a target protein joined to a ribosome binding site recognized and used for translation initiation in *E. coli* (Guarente *et al.*, "Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin," Cell 20: 543-553 (1980)) (Exhibit 2132). At page 543, through the first paragraph at page 544, this publication first summarizes previous approaches to the expression of proteins in their "native" state; *i.e.*, unfused to other proteins (Exhibit 2132, at 543,

Introduction, left col., first ¶, ll. 4-7). Guarente acknowledges that the previously disclosed Ptashne lab methods for expression of cloned genes in *E. coli* all had a very serious, fundamental limitation. No provision had been made for methods by which desired clones could be identified within the population of recombinant molecules produced according to the Ptashne lab methods, other than functional or immunological assays for the target protein. Exhibit 2132, at 544, left col., first full ¶. In actual practice, as noted by Guarente, identification of desired clones "may be laborious or impossible. (For example, affinity-purified antiserum to rabbit β-globin cross-reacts in a solid phase radioimmune assay with at least on *E. coli* protein)" (Exhibit 2132, at 544, left col., ll. 7-15).

104. The method disclosed by Guarente, which was intended to obviate this fundamental limitation to the actual use of the "Ptashne lab methods," is summarized in Fig. 1 (Exhibit 2132, at 545) and the first full paragraph, left column of p. 544. According to the method, the amino-terminal coding region of the target protein is fused, in frame, to an enzymatically-active carboxy-terminal coding sequence for the *E. coli* β-galactosidase. This hybrid target protein-β-galactosidase coding sequence is inserted into a plasmid with a unique restriction site upstream of the initiation codon of the target protein. To facilitate this step, three "host" plasmids are described, pLG200, pLG300, and pLG400. Each of these three plasmids carries a coding sequence for the enzymatically-active carboxy-terminal β-galactosidase coding sequence. However the reading frame for each pLG200, pLG300, and pLG400 is different so that insertion of a DNA fragment encoding at least the amino-terminal coding region of the target protein into each vector should guarantee that the desired hybrid protein will be expressed in one of the three.

105. According to the method of Guarente, a plasmid carrying the desired in-frame fusion protein can be opened at the unique restriction site, DNA sequences between that site and the ATG are removed (*e.g.*, by treatment with the combination of ExoIII and S1 or by treatment with Bal 31 alone), and a portable promoter, which carries a suitable SD sequence, inserted. The presence of the multiple copies of the *lac* operator within the portable promoter fragment titrates the low intracellular level of *lac* repressor and results in constitutive transcription from the *lac* promoter. The recombinant molecules are sealed with DNA ligase and used to transform a *lac*$^-$ *E. coli* host strain. Desired clones can be identified by colorimetric plate screens that distinguish *lac*$^-$ *E. coli* colonies from *lac*$^+$ *E. coli* colonies. Exhibit 2132, at 544, right col., first full ¶. These screens are based on the observation that successfully expressed protein fusions, in which the carboxy-terminal portion includes the indicated portion of β-galactosidase, are biochemically active and capable of degrading lactose. This activity results in a color change on MacConkey agar plates that contain lactose and the pH indicator, neutral red. Exhibit 2132, at 545, right col.

106. The methods used for the construction of plasmids pLG200, pLG300, and pLG400, are presented in Fig. 6 and Fig. 7 (Exhibit 2132, at 550 and at 551, left col.). The construction of these plasmids involved seven steps, which, as I understand them, are summarized below:

(a) A bacteriophage λ carrying the *E. coli* β-galactosidase gene (*lac* Z) was digested with KpnI and S1 (to generate a blunt-ended fragment) and then digested with EcoRI. The desired fragment carrying the carboxy-terminal coding portion of β-galactosidase gene was cloned between the EcoRI and PvuII (blunt) sites of the *E. coli* plasmid vector pBR322 to provide the recombinant plasmid pLG1.

(b) An EcoRI fragment was isolated from a plasmid designated pMC4. This fragment carried the gene for the *lac* repressor (*lac* I), the *lac* promoter and operator (*lac* po), and the amino-terminal coding sequence of the β-galactosidase gene. This EcoRI fragment was inserted into pLG1 to provide pLG2. Therefore, pLG2 carries the *lac* repressor, promoter, operator and an intact, biochemically active, β-galactosidase gene.

(c) For reasons that will become apparent in the subsequent steps, a *lac* Z^- derivative of pLG2 was isolated. This was accomplished by screening transformants carrying pLG2 on media containing a chromogenic substrate ("XG"), where *lac* Z^+ strains have an intense blue color and *lac* Z^- colonies are white. The desired *lac* Z^- mutant was one in which an insertion sequence disrupted the β-galactosidase gene.

(d) This *lac* Z^- derivative of pLG2 was introduced into an unidentified *E. coli* strain carrying an F'-factor, designated F'*lac*Δ71-56-14. I have been informed and it is my understanding that this F'-factor carries a genetic, in-frame fusion between the *lac* repressor gene (*lac*I) and the β-galactosidase gene (*lac* Z) such that the intervening *lac* po region was deleted. The resulting fusion protein possesses both *lac* repressor and β-galactosidase activity (*i.e.*, phenotypically *lac* Z^+). The intention was to allow natural *in vivo* recombination to occur between the homologous *lac* sequences of the *lac* Z^- derivative of pLG2 and those of F'*lac*Δ71-56-14, such that the *lac* po region of pLG2 would be lost.

(e) Plasmid DNA was isolated from the *E. coli* host carrying both F'*lac*Δ71-56-14 and pLG2 and used to transform a *lac* Z^- *E. coli* strain. The desired clones,

identified as intense blue colonies in the presence of XG carried the plasmid designated pLG5. This plasmid (pLG5) therefore encodes the *lac*I-*lac*Z fusion protein of F'*lac*Δ71-56-14; *i.e.*, the *lac* promoter naturally associated with *lac*Z has been removed.

(f) The goal of this step (step (f)) was to remove the promoter for the *lac*I gene along with the amino-terminal coding sequence for *lac*I as well. As noted in Fig. 6, the gene encoding the *lac*I-*lac*Z fusion protein carries five PvuII restriction sites. A partial digestion with PvuII was performed on pLG5 and a HindIII linker inserted. The desired plasmid (designated pLG200) would have a set of biochemical properties that could be identified in an appropriate *E. coli* host by virtue of growth and/or colorimetric signals generated on defined media, as described in the last paragraph of the legend to Fig. 7, at page 551 of Guarente.

(g) Plasmid pLG200 was digested with HindIII and the ends rendered blunt by incubation with a DNA polymerase and a 10-base pair BamHI linker inserted to provide pLG300. In turn, pLG300 was digested with BamHI, the ends filled in with DNA polymerase and a 10 base pair HindIII linker inserted to provide pLG400.

In view of the complexity of steps summarized above, it is my opinion that these plasmids and the method of Guarente were not within the scope of those of ordinary skill in the art as of March 19, 1980. Moreover, in view of the following paragraph, it would appear that the authors of the Guarente publication also believed their methods were new and not obvious over those previously disclosed in the art.

107. The newly developed method disclosed in Guarente (Ex. 2132) was the subject of UK patent application GB 2071671 A, which was filed March 16, 1981 (Exhibit 1078). The named inventors are Leonard P. Guarente, Mark Ptashne, and Thomas M. Roberts. The front page of this UK patent application indicates that it claims priority to an earlier U.S. patent application. I have been informed that the earlier U.S. application is U.S. patent application Serial No. 06/131,152, which was filed March 17, 1980 (Exhibit 1130).

VII. SUCCESSFUL EXPRESSION OF MATURE HUMAN FIBROBLAST INTERFERON IN *E. COLI*

A. Introduction

108. Despite all the uncertainties concerning whether any method would succeed in achieving biologically active non-glycosylated mature human fibroblast interferon in *E. coli*, three groups ultimately succeeded. Review of the publications reporting these successes illuminates the divergence of approaches and the difficulties that arose in making a DNA expressing the mature form of human fibroblast interferon. However, a theme common to all was the avoidance of brute force screening of a large populations of potential clones. This was accomplished by the use of either precise genetic engineering methods expected to yield a single recombinant construct or by the use of a new (June 1980), highly effective screening method of Guarente *et al.* (Exhibit 2132) that enabled the rapid and facile identification of desired clones from within the immense population of clones generated. By using the screening method disclosed, desired clones could be identified without reliance on an activity of the target protein nor even a physical property thereof, *e.g.*, the ability to bind an antibody.

109. The first published report on isolating a bacterial clone (TpIF319) thought to comprise DNA sequences complementary to mRNA encoding human fibroblast interferon was by Taniguchi *et al.*, "Construction and identification of a bacterial plasmid containing the human

fibroblast interferon sequence," Proc. Jpn. Acad. 55: 464-469 (1979) (Exhibit 2059). Subsequently, that research group used an internal PstI-BglII fragment of (TpIF319) to identify an apparently full length clone (TpIF319-13). Partial DNA sequence analysis of TpIF319-13 confirmed that this clone encoded at least the 34 amino-terminal amino acids of the precursor form of human fibroblast interferon, including the first 13 amino acids of mature human fibroblast interferon that corresponded to the previously published partial, amino-terminal amino acid sequence of human fibroblast interferon. Taniguchi *et al.*, "Molecular cloning of human fibroblast interferon cDNA," Proc. Nat'l Acad. Sci. USA 77: 4003-4006 (1980) (Exhibit 1014), at 4005, Fig. 4. The complete sequence of the precursor form of human fibroblast interferon was finally published in Taniguchi *et al.*, "The nucleotide sequence of human fibroblast interferon cDNA," Gene 10: 11-15 (1980) (Exhibit 1013), at 13, Fig. 2. The isolation and sequence of a cDNA encoding the precursor of human fibroblast interferon was independently and contemporaneously reported by Derynck *et al.*, "Isolation and structure of a human fibroblast interferon gene," Nature 285: 542-547 (1980) (Exhibit 1010), at 545, Fig. 4.

110. Although these papers (Taniguchi *et al.* (Exhibits 1013) and Derynck *et al.* (Exhibit 1010)) presented the deduced amino acid sequence of the mature form of human fibroblast interferon from these cDNAs encoding the precursor, neither described or even speculated as to any methods that could potentially be used (a) to tailor the cDNA to remove the DNA encoding the signal peptide from the DNA encoding the mature form, or (b) to properly position such tailored DNA with respect to bacterial expression control elements so as to be capable of direct expression of the mature form of human fibroblast interferon unaccompanied by its signal peptide or any portion thereof. It was in fact recognized by Taniguchi *et al.*, "Molecular cloning of human fibroblast interferon cDNA," Proc. Nat'l Acad. Sci. USA 77:

4003-4006 (1980) (Exhibit 1014), that "[i]t will be quite important to let the gene express correctly, because fused or incomplete proteins synthesized in bacteria may not be useful in clinical applications, even if such proteins exhibit antiviral properties." Exhibit 1014, at 4005, right col., third full ¶. The Fiers group also stated its intention to express the human fibroblast interferon gene in bacteria ("We are now in a position to reconstruct a plasmid with the total HF-IF genetic information under a prokaryotic transcription signal and to test for its expression in a bacterial system."). *See*, Derynck *et al.* (Exhibit 1010), at 546, right col., second full ¶. Again, there is no discussion in these papers of how to express the protein in bacteria.

111. At this time, there were two parallel approaches to heterologous expression of human fibroblast interferon in *E. coli.* In the first, the immediate goal was a critical "proof of concept" demonstration that biologically active recombinant fibroblast interferon could be produced in *E. coli.* For this demonstration, synthesis of a "polypeptide with interferon activity" would suffice, even if that protein were to include (or lack) one or more amino acids, or even include or exclude a substantial number of amino acids, provided the molecule exhibited any detectable level of interferon activity. In fact, it was subsequently reported that "fusion proteins" carrying three or even seven additional amino acids attached to the amino terminus of mature human fibroblast interferon possessed antiviral interferon activity (Itoh *et al.* (Exhibit 1120), at 157, "Abstract," second ¶, last line; at 159, right col., first full ¶, last seven lines, and at 161, left col., ll. 1-4). In contrast, the goal of the alternative approach was direct expression of a precisely defined protein molecule, mature human fibroblast interferon having a total of 165 or 166 amino acids (depending on whether *E. coli* removed the N-terminal methionine residue) unaccompanied by its signal peptide. The goal of the latter approach was to provide a clinically useful human therapeutic agent.

112. Construction of a DNA encoding a "polypeptide with interferon activity," was accomplished by assembly of a gene encoding a complex fusion protein comprising, *inter alia*, a portion of a prokaryotic protein joined to the 187-amino acid precursor form of human fibroblast interferon. In contrast, construction of a DNA encoding biologically active mature human fibroblast interferon having a total of 165 or 166 amino acids unaccompanied by its signal peptide (or any other amino acids) required tailoring of the mammalian gene. As demonstrated below, that was also a very complex endeavor, but it provided precisely defined, mature human fibroblast interferon.

113. In an intricate series of plasmid constructions, the Fiers group, of which I was a member, generated an expression plasmid which directed the expression of a fusion protein expected to contain the 98 amino-terminal amino acids of MS2 polymerase, a 27-amino acid "linker sequence" corresponding to the amino acid sequence encoded by the "non-coding region" upstream of the precursor, followed by the 187-amino acid precursor form of human fibroblast interferon. Derynck *et al.* (Exhibit 1011). I had considerable doubt that these constructions might express biologically active protein. Therefore, I was surprised when I was able to detect antiviral activity present in the extracts of heat-induced clones carrying the pPLc-HFIF-67-8 plasmid. Our subsequent publication of these results was the first to report recombinant expression of a "polypeptide having interferon activity." The "polypeptide with interferon activity" could have been generated by "haphazard proteolysis" of the expressed complex fusion protein either in *E. coli*, during sample preparation, or by processing or proteolysis by the human cells used for the antiviral CPE interferon assay. Exhibit 1011, at 197. Therefore, in view of the nature of the fusion protein encoded by the disclosed expression plasmid, and in the absence of any data to indicate otherwise, it could not be concluded that in fact, mature human fibroblast

interferon having a total of 165 or 166 amino acids and unaccompanied by its signal peptide had been produced. This is particularly apparent in hindsight, since it was subsequently reported that variants of mature human fibroblast interferon possessed antiviral interferon activity even though they carried additional amino acids attached to the amino terminus of the mature protein (*see* ¶ 111 above; and Itoh *et al.* (Exhibit 1120) at 157, "Abstract," second ¶, last line; at 159, right col., first full ¶, last seven lines, and at 161, left col., lines 1-4)).

B. Direct expression of mature human fibroblast interferon: Fiers

114. Following the publication of this critical "proof of concept" report (Derynck *et al.* (Exhibit 1011)), the Fiers group designed and carried out extensive and complex schemes of plasmid constructions, based on plasmid digestions and modifications using various enzymes. One of these approaches ultimately led to the successful, direct expression of mature human fibroblast interferon in *E. coli*, which was reported approximately three years later in Remaut *et al.*, "Inducible high level synthesis of mature human fibroblast interferon in *Escherichia coli*," Nucl. Acids Res. 11: 4677-4688 (1983) (Exhibit 1029). In the end the plasmid, which was reported to express biologically active mature human fibroblast interferon, was derived using an approach that should have yielded two juxtaposed initiator ATGs. The fact that the successful plasmid contained only one initiator ATG, as required for successful expression of the protein, is best explained by unanticipated overdigestion by exonuclease, thus illustrating the inherent problems and unpredictability of the approaches used.

115. Remaut states that, even as of 1983, and even in light of the large body of nucleotide sequence information regarding known ribosome binding sites, "our understanding of the specific factors that determine the efficiency of initiation at a given ribosome binding site is still very incomplete." Exhibit 1029, at 4677, first ¶. In contrast to other methods in which attempts were made to position a Shine-Dalgarno sequence at various distances from an ATG

initiation codon, the approach disclosed in Remaut, in essence, attached the coding sequence of the target protein to a pre-formed ribosome binding site, *i.e.*, one in which a promoter, Shine-Dalgarno sequence and ATG had already been demonstrated to induce high-level expression of a different protein. In particular, Remaut describes the construction of an expression vector, pPLc245, which includes the strong phage promoter ("λ P_L promoter"), and an efficient combination of Shine-Dalgarno sequence and ATG initiation codon, which is followed by a "polylinker" sequence for insertion of the coding sequence of a target gene. Exhibit 1029, at 4680, Fig. 1; at 4679-4681, Section 1 "Construction of pPLc245"; and at 4686, first full ¶). The λ P_L promoter can be tightly regulated by the λ repressor protein expressed in the host *E. coli*. Specific *E. coli* host cells employed by Remaut (*see* Exhibit 1029, at 4678, Table 1) encode a temperature-sensitive λ repressor protein ("cI857"), such that transcription from the λ P_L promoter is tightly suppressed at 28 °C but fully induced at 42 °C. The coding sequence for the mature form of human fibroblast interferon was obtained from a plasmid designated pPLcRX19 in which the ATG initiation codon had been fused to the DNA sequence (TCTAGA) recognized by the restriction endonuclease XbaI (*i.e.*, --TCTAG**ATG**--). Construction of this key intermediate plasmid, having the precisely-placed unique restriction site, is described in the (1983) Ph.D. thesis of P. Stanssens of Dr. Fiers' laboratory. Digestion of pPLcRX19 with XbaI, followed by treatment with S1 nuclease should provide a blunt-ended DNA molecule starting with the ATG initiation codon of mature fibroblast interferon. Similarly, digestion of the expression vector pPLc245 with SalI, followed by treatment with nuclease S1 should have provided a blunt ended DNA ending with ATG. Joining these two fragments was predicted to generate a molecule carrying the λ P_L promoter and a SD sequence followed by two ATG codons. Exhibit 1029, at 4682, ll. 21-24. However, only one ATG codon was found, apparently

the result of S1 "nibbling" of one or both of the ends of the molecules to be joined. Exhibit 1029, at 4681-4682, Section 2 "Insertion of the human β-interferon gene into pPLc245." Identification of the desired clones that expressed human β-interferon was facilitated by the authors' observation that even "marginal expression" of this protein resulted in a dramatic decrease in the growth rate of the host *E. coli*. Transformants therefore were first screened for those growing well at 28 ºC but poorly at 42 ºC (Exhibit 1029, at 4682, first ¶; at 4684, first full ¶ and Fig. 3). The authors noted that expression of mature human fibroblast interferon in their λ P_L system (*e.g.*, from pPLc245-HIF 25) was substantially higher than that reported previously with other *E. coli* expression systems (Exhibit 1029, at 4686, second full ¶). With respect to the apparently toxic (to *E. coli*) consequences of expression of human fibroblast interferon, the authors stated:

> this observation stresses the importance of being able to regulate the expression of a cloned gene. We observe that a bacterial strain containing pPLc245-HFIF25 grows quite normally at 28 ºC under non-expressing conditions, but is severely impaired in its growth rate at 42 ºC when the inserted gene is maximally expressed. As a consequence, it would appear that *a host-vector system which does not allow full control over the expression of an inserted gene will tend to be unstable and will have negative selection value.* Exhibit 1029, at 4687, ll. 16-23.

116. It is my opinion that these observations and results validate earlier concerns in the art that it may be difficult, or even impossible, to express a mammalian protein in *E. coli* if that target protein is toxic to the host cell.

C. Direct expression of mature human fibroblast interferon: Goeddel

117. The genetic engineering methods developed by Goeddel (Exhibit 1012) for direct expression of the mature human fibroblast interferon provided two DNA molecules encoding the 166-amino acid mature form of human fibroblast interferon unaccompanied by its signal peptide. The molecules differed only in that one contained an extra base pair in order to re-create an EcoRI restriction site upon insertion into a *lac* promoter-based expression plasmid. According to

the method of Goeddel, a 1200-base pair HhaI double-stranded DNA restriction fragment carrying the entire coding sequence for the 187-amino acid precursor form of human fibroblast interferon was denatured and annealed with either a first primer (**ATG**AGCTACAAC) or a second primer (C**ATG**AGCTACAAC) (Exhibit 1012, at 4065, ll. 1-2). These primers encompass the first four codons (for methionine, serine, tyrosine, and asparagine, respectively) of the mature form of human fibroblast interferon. Although Goeddel discloses the use of both primers and the construction of two expression systems, for clarity, only the use of the second primer is discussed here. The subsequent step was a "primer extension reaction" in which *E. coli* DNA polymerase (Klenow fragment) extended the oligonucleotide primer while copying the template DNA. This enzyme also removed the remaining single stranded "tails" (*i.e.*, unpaired single stranded template DNA upstream of the amino-terminal methionine codon of mature human fibroblast interferon; *see* Goeddel *et al.* (Exhibit 1012), at 4066, Fig. 4). Digestion of the products generated in the primer extension reaction with PstI released the amino-terminal coding sequence of the mature form of human fibroblast interferon unaccompanied by the coding sequence of human fibroblast interferon signal peptide. This fragment was then enzymatically joined to a PstI-BglII DNA fragment encoding the carboxy-terminal coding sequence of human fibroblast interferon. Gel electrophoretic separation of the products of that enzymatic joining provided a single, defined, 505-base pair duplex DNA molecule encoding the 166-amino acid mature form of human fibroblast interferon unaccompanied by the coding sequence of the signal peptide. This 505-base pair duplex DNA was then enzymatically inserted into a *lac* promoter-based expression vector as depicted in Fig. 6 and an *E. coli* host transformed with the resulting recombinant molecules (Exhibit 1012, at 4068, Fig. 6). Since correctly assembled expression plasmids would create an EcoRI site, DNA was isolated from a number of random

transformants and screened for the presence of this diagnostic restriction site. As expected, the majority of isolates gave the predicted restriction enzyme digestion patterns (Exhibit 1012, at 4069, first full ¶). Moreover, of 12 clones identified as having the predicted EcoRI site, 11 exhibited equivalent levels of interferon activity (Exhibit 1012, at 4069, third ¶).

118. The method used by Goeddel for expression of mature human fibroblast interferon (*see* ¶ 117) was not the same method used by Goeddel for expression of mature human growth hormone (*see* Exhibit 1057). Moreover, Goeddel's method (*see e.g.*, Exhibit 1012, at 4066, Fig. 4) could have been successfully carried out using: (1) the knowledge of the nucleotide coding sequence for, *e.g.* the first four amino acids of the mature protein; (2) a DNA fragment identified as carrying the entire coding sequence for mature human fibroblast interferon (identified and isolated as described in Goeddel *et al.* (Exhibit 1012) and, *e.g.*, identified as carrying the coding sequence for the first 4 amino-terminal amino acids and large enough to encompass the coding sequence of a protein of the molecular weight of human fibroblast interferon); and (3) a basic restriction map identifying *inter alia*, a restriction site beyond the end of the coding sequence. That is, Goeddel did not need to know the entire coding sequence to be able to express human fibroblast interferon.

D. Direct expression of mature human fibroblast interferon: Sugano

119. The cDNA cloned by Sugano was used in a collaboration with another research group, the Ptashne group, to express human fibroblast interferon in bacteria. The results of these efforts that led to the direct expression of what was purported to be the mature fibroblast interferon were reported in October 1980. Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009). This publication is discussed in detail in the following paragraphs.

120. Using a series of elegant and complex plasmid constructions, the Taniguchi/Ptashne group generated an expression plasmid reported to express the mature form of human fibroblast interferon. A critical aspect of these plasmid constructions was the use of Bal 31 exonuclease to successively remove nucleotides from the DNA segment encoding the precursor form in order to eliminate the entire signal peptide. Such sequential digestion results in a large number of different DNA molecules terminating at various positions ranging from upstream of the signal peptide to within the coding region. To isolate a molecule free of DNA encoding the signal peptide, the Taniguchi/Ptashne group used a powerful screening system that allows identification of those constructs, based on the generation of fusion proteins with the amino-terminus of the human fibroblast interferon fused to the carboxy-terminus of β-galactosidase. The biochemically active β-galactosidase fusion protein produced from that construct could be readily identified using a colorimetric plate screen (Exhibit 1009, at 5230, left col., second full ¶ under "Results," to right col., first 3 lines). The construct was hoped to have no remaining DNA portion encoding the signal peptide and the ATG of the mature protein properly positioned with respect to bacterial expression control elements. The combination of exonuclease digestion and screening techniques succeeded, and resulted in the generation of an expression plasmid capable of making what was purported to be mature human fibroblast interferon. In my opinion, prior to their actual success, it could not have been predicted with a reasonable degree of certainty that this approach would actually work. It should also be noted that although Taniguchi cloned the cDNA encoding the precursor form of human fibroblast interferon in his lab in Japan, he did not successfully pursue the expression of human fibroblast interferon in his own lab; rather he sought collaboration with the Ptashne lab in the United States. Weissmann, "The Cloning of Interferon and Other Mistakes," IN INTERFERON 3,

Academic Press Inc. (New York, NY), 1981, pp. 101-134 (Exhibit 1131), at 119-120. This illustrates the challenges and non-obviousness imposed by the attempts to express human fibroblast interferon in bacteria. The complexity and value of the screening method used by Taniguchi and the members of Ptashne's lab are discussed in greater detail in the following paragraphs.

121. Taniguchi *et al.*, "Expression of human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009). This publication discloses the application of the new method of Guarente (published in June 1980) to the problem of expression of the mature form of human fibroblast interferon. In the first stage of the experiments the authors constructed a plasmid in which the amino-terminal coding region of the precursor form of human fibroblast interferon was fused to the I'-β-galactosidase "reporter" protein of pLG300. Exhibit 1009, at 5231, Fig. 1 and at 5230, right col., first full ¶. The desired plasmid carrying this fusion was constructed in two steps.

122. In the first step, four separate DNA molecules were combined to provide the plasmid designated pLG111, which carried, *inter alia*, the complete gene coding for the precursor form of human fibroblast interferon as well as the coding sequence for the I'-β-galactosidase "reporter" protein of pLG300. Exhibit 1009, at 5230, right col., ll. 12-23. In addition, pLG111 was constructed in such a manner that a unique HindIII site was created 3 base pairs upstream of the ATG initiation codon for the precursor form of human fibroblast interferon.

123. In the second step, three DNA fragments were prepared and combined. Exhibit 1009, at 5230, right col., ll. 23-35. In particular, a PstI-PstI fragment of pLG111 was isolated, which included the amino-terminal coding sequences for the precursor form of human fibroblast interferon as well as the amino-terminal coding region of the ampicillin-resistance gene (amp^r).

A second fragment was prepared and isolated which encompassed a portion of the coding sequence of human fibroblast interferon that is found between the (unique) PstI site in that gene and the nearest, downstream HinfI site (which HinfI site apparently had been rendered flush-ended by DNA polymerase I). The third fragment isolated was a PstI-BamHI fragment (which BamHI site had also been rendered flush-ended by DNA Polymerase I) comprising the coding sequence of the carboxy-terminal half of the amp^r gene and the I'-β-galactosidase "reporter" protein of pLG300. The resulting molecule, designated pTR56, carries a unique HindIII site 3 base pairs upstream of the ATG initiation codon for that amino-terminal coding region of the precursor form of human fibroblast interferon fused to the I'-β-galactosidase "reporter" protein. In addition, pTR56 is indicated as having a unique BamHI site further "upstream" from that ATG initiation codon.

124. In a first set of experiments, pTR56 was linearized by digestion with HindIII and then digested with nuclease Bal31 (rather than the combination of ExoIII followed by S1) and samples taken at various time points. The DNA digested in this manner was then cleaved with BamHI and a BamHI-PvuII fragment (the PvuII end is a is flush end) carrying a *lac* promoter, SD sequence and 5 additional base pairs, was inserted. This recombinant DNA was used to transform an *E. coli* host strain that lacks β-galactosidase activity. Transformants were selected for resistance to ampicillin and screened on a colorimetric indicator medium to identify clones expressing the human fibroblast interferon precursor-β-galactosidase fusion protein. Notably, despite the close proximity of the unique HindIII site to the ATG initiation codon (6 base pairs), only 5% of the transformants were identified as lactose-utilizing (Exhibit 1009, at 5232, left col., first full ¶), indicating successful expression of the human fibroblast interferon precursor - I'-β-galactosidase fusion protein. The one clone identified and characterized

(pLG104) was found to have a seven-base pair spacing between the *lac* SD and the ATG initiation codon of the precursor form of human fibroblast interferon. Thus, the Bal31 treatment had removed approximately 4 base pairs. Subsequent manipulations replaced the β-galactosidase portion of the fusion protein with the carboxy-terminal coding sequences of the precursor form of human fibroblast interferon to provide the plasmid pLG104R. This plasmid reportedly expressed the precursor form of human fibroblast interferon, which was found to be (1) totally devoid of interferon activity and (2) extremely unstable, even as compared to the mature form of the same protein (Exhibit 1009, at 5233, right col., first ¶) and (3) was not processed to the mature form (Exhibit 1009, at 5232, right col., first ¶, last sentence).

125. In a second set of experiments, pTR56 was linearized by digestion with HindIII and then digested for more extended periods of time with nuclease Bal31 (rather than the combination of ExoIII followed by S1) in order to remove most or all of the human fibroblast interferon signal peptide, which in pTR56 would be approximately 70 base pairs. The DNA digested in this manner was then cleaved with BamHI and a BamHI-PvuII fragment (the PvuII end is a flush end) carrying a *lac* promoter, SD sequence and 5 additional base pairs, was inserted. The recombinant DNA was used to transform an *E. coli* host strain that lacks β-galactosidase activity. Transformants were selected for resistance to ampicillin and screened on a colorimetric indicator medium to identify clones expressing the human fibroblast interferon precursor-β-galactosidase fusion protein. Notably, 99.99% of the transformants were not lactose-utilizing colonies; *i.e.*, only 0.01% were lactose-utilizing (Exhibit 1009, at 5232, left col., first full ¶). The one clone identified and characterized (pLG117) was also found to have a seven-base pair spacing between the *lac* SD and the ATG initiation codon the mature form of human fibroblast interferon. As noted, this clone was extremely rare, *i.e.*, 1 in 10,000. This

screen yielded another clone, designated pLG115. It was determined that even though this plasmid still possessed the ATG initiation codon for the mature form of human fibroblast interferon, translation nonetheless initiated with the ATG codon at position 175 of the DNA sequence of Taniguchi *et al.* (Exhibit 1009, at 5232; left col., footnote). Thus, removal of part of the coding sequence of the human fibroblast interferon signal peptide somehow enhances translation from the internal ATG codon, bypassing the ATG coding for the amino-terminal methionine residue of mature human fibroblast interferon. In view of this result and other considerations, discussed above, it is my opinion that it is apparent that simply removing the ATG initiation codon of the coding sequence for the precursor form of human fibroblast interferon does not guarantee that the next ATG codon, which happens to be that for the amino-terminal amino acid of the mature form, will be used for initiation of translation.

126. Subsequent manipulations replaced the β-galactosidase portion of the fusion protein of pLG117 with the carboxy-terminal coding sequences of the human fibroblast interferon to provide the plasmid pLG117R. This plasmid was purported to express the mature form of human fibroblast interferon, which was found to exhibit only 1% of the expected level of interferon activity based upon the amount of protein synthesized. By way of explanation for this observation, the authors suggested, *inter alia*, that perhaps unglycosylated, bacterially synthesized human fibroblast interferon has a low specific activity in their *in vitro* assay (Exhibit 1009, at 5233, left col., last ¶).

E. Subsequent Reports of Direct Expression of Mature Human Fibroblast Interferon by Taniguchi

127. In addition to the above, Dr. Taniguchi is a co-author or co-inventor of at least four other reports directed toward heterologous expression of mature human fibroblast interferon in *E. coli*. These are, in chronological order, (1) European Patent Application No. 0064681 A2

(claiming priority to April 30, 1981) (Exhibit 1121), (2) U.S. Patent No. 4,686,191 (claiming priority to December 25, 1981) (Exhibit 1120), (3) Taniguchi *et al.*, *IN* "Proceedings of the IVth International Symposium on Genetics of Industrial Microorganisms (1982)" (Y. Ikeda & T. Beppu, eds.; Kobansha Ltd., Tokyo; 1983) (Exhibit 1068), and (4) Itoh *et al.*, "Efficient expression in *Escherichia coli* of a mature and a modified human interferon-β_1," DNA 3: 157-165 (1984) (Exhibit 1119).

128. European Patent Application No. 0064681 A2 (Exhibit 1121) discloses the construction of an expression plasmid designated pTuIFN β-5 (Exhibit 1121, at 20, Fig. 3), in which a DNA encoding mature human fibroblast interferon having a total of 166 amino acids was placed under the expression regulatory elements derived from the *E. coli* gene for the protein chain elongation factor EF-Tu. Salient features of the construction of pTuIFN β-5 include the following. Plasmid pTR56 (*see* Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980) (Exhibit 1009)), was opened with the restriction enzyme HindIII, digested with Bal31, and, *inter alia*, a BamHI linker (CCGGATCCGG) inserted (Exhibit 1121, p. 5, ll 6-15). After an additional step, the resulting plasmid, designated pTuIFN β-4, lacked the coding sequence for the first six amino acids of the mature form of human fibroblast interferon but now possessed a newly-created HpaII site as a result of the insertion of the BamHI linker (Exhibit 1121, p. 5, ll 16-30). Subsequently, pTuIFN β-4 was digested with HpaII and joined to a synthetic DNA molecule that encoded the first six amino acids of the 166-amino acid mature form of human fibroblast interferon, and the Tu promoter to provide pTuIFN β-5 (Exhibit 1121, p. 6, l. 25 to p. 7, l. 12). A noteworthy feature of pTuIFN β-5, exploited in the following publications, is the presence of a Cla1 restriction site (ATCGAT) that overlaps the ATG initiation codon for mature human

fibroblast interferon, *i.e.*, ATCG**ATG** (Exhibit 1121, p. 8, ll 11-16). An *E. coli* strain transformed with pTuIFN β-5 was alleged to produce almost ten times as much human fibroblast interferon as did a strain transformed with pLT117R (Exhibit 1121, p. 15, ll 1-12).

129. U.S. Patent No. 4,686,191, "Recombinant Plasmid Containing Human Interferon-Beta Gene," (Exhibit 1120), describes the construction of a number of expression plasmids employing expression elements derived from the *E. coli trp* operon which produce substantial amounts of human fibroblast interferon (*see e.g.*, Exhibit 1120, at col. 8, Table 1). In essence, these expression systems were constructed by replacing the Tu regulatory elements of pTuIFN β-5 with those derived from genes of the *trp* operon of *E. coli* (*see e.g.*, Exhibit 1120, at sheet 8 of 10, Fig. 8).

130. Taniguchi *et al.*, "Expression of the Human Interferon-β_1 Gene in Heterologous Host Cells," *IN* "Proceedings of the IVth International Symposium on Genetics of Industrial Microorganisms (1982)" (Y. Ikeda & T. Beppu, eds.; Kobansha Ltd., Tokyo; 1983) (Exhibit 1068). This publication also reports the construction of expression plasmids in which regulatory elements of the genes of the *trp* operon of *E. coli* were joined to a coding sequence for mature human fibroblast interferon. Again, these expression plasmids were constructed by replacing the Tu regulatory elements of pTuIFN β-5 with those derived from genes of the *trp* operon of *E. coli* (*see e.g.*, Exhibit 1068, at 205, Fig. 4).

131. Itoh *et al.*, "Efficient expression in *Escherichia coli* of a mature and a modified human interferon-β_1," DNA 3: 157-165 (1984) (Exhibit 1119). This publication describes the construction of additional expression plasmids in which regulatory elements of the *E. coli trp* operon genes and lipoprotein gene were used to drive expression of not only mature human fibroblast interferon but also variants thereof carrying additional amino-terminal amino acids

(*see* Exhibit 1119, at 164, Fig. 4). This publication also disclosed the varying levels of human fibroblast interferon production as a function of the spacing between the SD sequence and the ATG initiation codon in a number of different expression systems (*see* Exhibit 1119, at 162, Table 1 and at 163, Table 2).

F. Summary and Conclusions

132. Stepping back from the technical details of the publications by these three groups, I note the following: (a) none of these expression strategies used by the Fiers and the Taniguchi/Ptashne groups applied the approach described in the Goeddel hGH paper; (b) both groups used divergent approaches that included a lot of trial and error, and none of these was based on the use of synthetic DNA, illustrating that, at that time, no defined, general method to express eukaryotic proteins in bacteria was known to and agreed upon by the investigators; (c) the cloning of cDNA encoding the precursor form of human fibroblast interferon and the expression of mature human fibroblast interferon were considered not only by both the Taniguchi/Ptashne group and the Fiers group, but also by the reviewers and editors of these top journals, as separate and significant developments along the way to expressing mature human fibroblast interferon, that merited separate, highly visible publications. Indeed, the Nature editors and the Fiers group considered their expression of human fibroblast interferon as a fusion protein sufficiently meritorious to warrant separate publication prior to the ultimate success in expressing the mature form; and (d) in fact, Goeddel's own publication of his success in expressing the mature form also did not apply the method published in the Goeddel hGH paper (see, Goeddel *et al.*, "Synthesis of human fibroblast interferon by *E. coli*," Nucl. Acids Research 8: 4057-4074 (1980) (Exhibit 1012)).

133. In my opinion, the above facts clearly demonstrate that the researchers of the Fiers group, the Taniguchi/Ptashne group and the Goeddel group, and the reviewers and editors

of the journals in which the papers were published, contemporaneously believed that successful direct expression of the mature form of human fibroblast interferon was a significant accomplishment above and beyond the cDNA cloning of the precursor gene. That is, this work was deemed not to be "routine" by the editors and reviewers of these journals. This conclusion is also supported by subsequent publications and patent applications from the Sugano and Taniguchi laboratories disclosing additional approaches to the expression of mature human fibroblast interferon (*see* ¶¶ 127-131). These facts also demonstrate that these accomplishments were far beyond the skill of the ordinary researchers at that time, and beyond the knowledge available in the literature at that time. That significant accomplishment could only have been achieved by extensive experimentation beyond the level of skill in the art at the time.

134. All of these research groups acted upon their independent recognition that they needed to develop and apply techniques to obtain DNA encoding a polypeptide comprising the 166-amino acid mature form of human fibroblast interferon, in which such DNA is entirely free of the DNA encoding the signal peptide found in the naturally occurring gene.

135. Moreover, all three of these research groups further recognized at the time that only by constructing an expression plasmid in which the DNA encoding mature human fibroblast interferon, having a total of 166 amino acids and unaccompanied by its signal peptide, properly positioned with respect to bacterial expression control elements could direct expression of the mature form be achieved in bacteria.

VIII. JAPAN PATENT APPLICATION NO. 33931/1980

136. I have been asked to consider whether, given Japan Application Serial No. 33931/80 ("the Japan '931 application"), filed March 19, 1980 (Exhibit 2012), one skilled in the art, as of that filing date, would have been able to construct a DNA of encoding mature human

fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide, without undue experimentation.

137. I have been advised that in order for a patent application to describe an invention, the inventor's disclosure must be sufficient to convince one of ordinary skill in the art that the inventor actually possessed the invention, *i.e.*, the inventor actually invented the claimed invention.

138. Based upon my review of the translation of the Japan '931 application in English (Exhibit 2013), it is my opinion that the Japan '931 application would not have enabled a person skilled in the art as of March 19, 1980 to make DNA encoding mature human fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide, without undue experimentation, as further detailed below.

139. Similarly, it is my opinion that the Japan '931 application would not have enabled a person skilled in the art as of March 19, 1980 to make a microbial replicable expression vector capable of expressing mature human fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide, without undue experimentation, as further detailed below.

140. To successfully produce a mature human fibroblast interferon free of its signal peptide, two key elements must be achieved: (a) DNA encoding the precursor form must be "tailored," or modified, to remove those codons encoding the entire signal peptide and (b) the ATG corresponding to the first amino acid of the mature form must be properly positioned with respect to bacterial expression control elements.

141. Nowhere in the Japan '931 application does Sugano describe a method for tailoring the human fibroblast interferon gene, which encodes the amino acids of the signal

peptide, to remove those codons encoding the entire signal peptide. Nor does the Japan '931 application describe a method for making an expression vector in which the ATG corresponding to the first amino acid of the mature form is properly positioned with respect to bacterial expression control elements such that it is capable of directing expression of a DNA encoding a mature human fibroblast interferon having a total of 166 amino acids and free of the fibroblast interferon signal peptide.

142. The naturally-occurring gene for human fibroblast interferon codes for the "precursor form," also called "pre-interferon." Specifically, the gene codes for a pre-interferon having a total of 187 (21 + 166) amino acids. The first (*i.e.*, amino-terminal) 21 amino acids of the pre-interferon constitute a "presequence," or, as used herein, a "signal peptide." The following (*i.e.*, carboxy-terminal) 166 amino acids of the pre-interferon constitute mature fibroblast interferon. In mammalian cells expressing the naturally occurring gene, the 187-amino acid pre-interferon is initially produced and this pre-interferon is transported through the endoplasmic reticulum of the mammalian cell, where the 21-amino acid signal peptide is removed, and the 166-amino acid mature fibroblast interferon is then secreted from the cell. (Houghton *et al.*, "The absence of introns within a human fibroblast interferon gene," Nucl. Acids Res. 9: 247-266 (1981) (Exhibit 1067), at 259, first full ¶, l. 6 through 260, l. 6; Taniguchi *et al.* (Exhibit 1068) at 205, ll. 3-5, citing Taniguchi *et al.* (Exhibit 1013); Taniguchi *et al.* (Exhibit 1013) at 13, Fig. 2; and Knight, "Human fibroblast interferon: Amino acid sequence analysis and amino terminal amino acid sequence," Science 207: 525-526 (1980) (Exhibit 1037), at 526, center col., first ¶ (13 amino-terminal amino acids of human fibroblast interferon)). Naturally-occurring mature fibroblast interferon (*i.e.*, produced by mammalian cells) is glycosylated (Knight, "Interferon: Purification and initial characterization from human diploid

cells," Proc. Nat'l Acad. Sci. USA 73: 520-523 (1976) (Exhibit 1036), at 520, Abstract; at 522, left col. fourth ¶ through the first ¶ in the right col.) whereas mature fibroblast interferon produced in bacterial cells is nonglycosylated.

143. To successfully produce a mature human fibroblast interferon free of its signal peptide, two key elements must be achieved: (a) DNA encoding the precursor form must be "tailored," or modified, to remove those codons encoding the entire signal peptide and (b) the ATG corresponding to the first amino acid of the mature form must be properly positioned with respect to bacterial expression control elements.

144. As of March 19, 1980, it was known that mammalian mRNAs differed considerably from bacterial mRNAs (*see e.g.* Steitz (Exhibit 1109), at 489, last ¶, which ends on 490 and at 490, last ¶, which ends on page 491). For example, bacterial mRNAs naturally contain a sequence, known as the "Shine-Dalgarno sequence," in the 5' untranslated leader region of the mRNA, which is required for efficient bacterial ribosome binding and initiation of translation. Shine and Dalgarno, "Determinant of cistron specificity in bacterial ribosomes," Nature 254: 34-38 (1975) (Exhibit 1043). Thus, to achieve efficient translation in bacteria, mRNAs comprising a mammalian protein coding sequence had to be engineered to contain a Shine-Dalgarno sequence in a proper sequence context.

145. As of March 19, 1980 the art had demonstrated that:

(a) heterologous expression of a coding sequence in *E. coli* required a promoter operatively associated with that gene that would direct transcription of the coding sequence to provide mRNA;

(b) heterologous expression of a coding sequence in *E. coli* required the presence of a "ribosome binding site" (which includes the "Shine Dalgarno" sequence and the

ATG initiation codon) disposed within the mRNA, that would direct initiation of translation of mRNA encoding the gene product; and

(c) heterologous expression of a coding sequence in *E. coli* depended upon appropriate positioning and spacing of each of the promoter, the ribosome binding site, and the initiation codon relative to one another.

146. Applying the principles of paragraph 145, and starting with a DNA molecule encoding the 187-amino acid precursor form of human fibroblast interferon expression of the mature form of human fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide in *E. coli* required removal of the DNA encoding the human fibroblast interferon signal peptide so that the ATG initiation codon of the mature form of human fibroblast interferon could be properly positioned relative to a promoter and a ribosome binding site.

147. Therefore, in view of the above, as of March 19, 1980 the nascent fields of recombinant DNA and heterologous gene expression were not sufficiently developed to the extent that one of ordinary skill in the art would have had a reasonable expectation of being able to successfully carry out the "tailoring" required to generate a DNA encoding the mature form of human fibroblast interferon, unaccompanied by the human fibroblast interferon signal peptide, and then to express that encoded protein in *E. coli*, absent a specific and detailed description of methods for doing so.

148. As of March 19, 1980, one of ordinary skill in the art would have understood that an *E. coli* strain identified as "capable" of expression of the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by a signal peptide would be an organism carrying a properly tailored and engineered gene directly encoding that protein, *e.g.*, as

described in paragraphs 144-6 above, that actually directed synthesis of the mature form of human fibroblast interferon.

149. The section of the Japan '931 patent including the last paragraph on page 2, and extending through the third full paragraph at page 7, presents a generic description of a method for isolating interferon coding sequences. According to the method presented, mRNA is isolated from poly(I):poly(C) induced cells, cDNA is prepared from the mRNA and inserted into an *E. coli* vector using methods and reagents known in the art. Recombinant plasmids carrying an insert complementary to human interferon mRNA are selected using unspecified procedures. The selected molecules are used as probes to identify other recombinant plasmids carrying an insert that codes for a polypeptide with interferon activity, again according to methods not specified in this section.

150. The "Brief Description of the Drawing" of the Japan '931 application states that the Figure depicts the restriction map of two clones, designated #319 and #319-13 (Exhibit 2013, at 16, third ¶). It also states that former was used to make the latter. Both inserts are alleged to show complementarity to human fibroblast interferon mRNA (Exhibit 2013, at 2, ll. 13-23).

151. The section of the Japan '931 application from page 3-7 describes methods known in the art for inducing interferon mRNA production by human cell cells, conversion of the mRNA to cDNA, and cloning the resulting double stranded cDNA fragments (Exhibit 2013, at 3, second full ¶ to 7, second full ¶). In particular, this section describes size fractionation of induced mRNA and the use of frog oocytes for detecting interferon mRNA per se as well as cloned DNA molecules capable of hybridizing to interferon mRNA.

152. At page 7, the Japan '931 application alleges that the plasmids disclosed would be useful for mass production of human fibroblast interferon in *E. coli* or eukaryotic cells (Exhibit 2013, at 7, third full ¶). Methods to be used to achieve these ends are not described.

153. In the section spanning pages 7 and 8, the Japan '931 application describes priming and superinduction of human fibroblasts according to a published procedure for stimulation of interferon mRNA synthesis (Exhibit 2013, at 7, last full ¶ to 8, Table 1). Total RNA was isolated and the polyA-containing mRNA fraction was isolated and size fractionated on a sucrose gradient generally according to methods known in the art. The presence of interferon encoding mRNA in gradient fractions was identified by translation in frog oocytes followed by an *in vitro* assay for antiviral activity.

154. In the section spanning the first paragraph at pages 9 and 10, the Japan '931 application indicates that the mRNA identified (as described in the previous paragraph), was converted to a cDNA and then to a double stranded DNA molecule carrying a protruding single-stranded tail of "T" residues (Exhibit 2013, at 9, first full ¶ through 10, second full ¶). Vector DNA, the *E. coli* plasmid pBR322, was digested with EcoRI and the protruding 5' ends removed with λ exonuclease. The thus exposed 3' ends were modified by addition of a single stranded tail of "A" residues. The "T-tailed" cDNA and the "A-tailed" plasmid were annealed, used to transform the *E. coli* host strain, χ1776, and ampicillin resistant transformants were selected. The methods employed in this section were generally known according to those familiar with the art.

155. The Japan '931 application then describes the preparation of two radioactively labeled probes that were then used to screen the ampicillin-resistant transformants obtained (above) to identify those carrying cloned inserts complementary to interferon mRNA. The

probes were prepared by labeling mRNA isolated either from uninduced cells (probe "A") or induced cells (probe "B"). Four transformants were identified that hybridized preferentially to probe B (Exhibit 2013, at 10, third full ¶ through 11, first full ¶ and Table 2). Each was tested for its ability to hybridize to interferon encoding mRNA by the methods disclosed (Exhibit 2013, at 11, second full ¶ through 12, first full ¶ and Table 3). The positive results with one clone, designated #319 were confirmed (Exhibit 2013, at 12, last ¶ through 13, second ¶ and Table 4) and a restriction map of that plasmid generated as described (Exhibit 2013, at 13, second and third full ¶¶).

156. The Japan '931 application continues, describing the isolation of a PstI-BglII fragment of the insert carried by isolate #319. This fragment was radioactively labeled and used as a probe to re-screen the initial population of ampicillin-resistant transformants. The recombinant plasmid carrying the largest complementary insert was designated #319-13. A restriction map of recombinant plasmid #319-13 was generated and the insert sequenced according to published methods (Exhibit 2013, at 13, last ¶ through the end of 14). The DNA sequence obtained is presented (Exhibit 2013, at 15).

157. The Japan '931 application notes the importance of the finding of the 13-amino acid sequence reported by Knight (Exhibit 1037) (Exhibit 2013, at 16, ll. 1-7) and then suggests that "the plasmid encompasses the entire coding region of the protein of the above mRNA and probably the coding region of the signal peptide" (Exhibit 2013, at 16, ll. 8-11). However, the Japan '931 application does not identify the reported 187-amino acid sequence as a precursor protein, nor does this application identify either the 166-amino acid mature form of human fibroblast nor the 21-amino acid signal peptide. Notwithstanding this fact, in the second paragraph at page 16, the Japan '931 application states that:

> The supports that transformation of the plasmid or mRNA inserted therein to other expression plasmids enables a host such as Escherichia coli to produce interferon. Exhibit 2013, at 16, second ¶.

158. In contrast to the assertions of the above quoted paragraph, the specification of the Japan '931 application does not provide any method or even a reference to a method by which the reported 777 bp DNA sequence at page 15 might be tailored to provide a DNA encoding mature human fibroblast interferon, having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide, much less a method for expressing the mature protein encoded by such a tailored DNA

159. The specification of the Japan '931 application provides the recombinant molecule #319-13, as well as the DNA sequence encoding the 187-amino acid precursor form of human fibroblast interferon. In view of the 13 amino-terminal amino acids of human fibroblast interferon disclosed by Knight (Exhibit 1037), those of ordinary skill in the art should have been able to envision a DNA encoding mature human fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide (*e.g.*, a DNA consisting of base pairs 70-567 of the sequence provided at p. 15) (Exhibit 2013, at 15). Sugano's specification, however, did not do so.

160. Those skilled in the art should have recognized that the two plasmids disclosed in the Japan '931 application, #319 and #319-13, would not be able themselves to express mature, biologically active human fibroblast interferon following their introduction into a suitable host, for the reasons provided in the following paragraphs.

161. Inspection of the restriction map for plasmid #319 (Fig. 1(a)) indicates that plasmid #319 is missing the DNA coding sequence for the presequence through and beyond the amino terminus of the mature protein (Exhibit 2013, at 17, Fig. 1(a)). Plasmid #319 apparently does not include the full 166-amino acid sequence of mature human fibroblast interferon.

(Taniguchi *et al.*, (Exhibit 1014), at 4004, Fig. 2, right col., last ¶). The specification of the Japan '931 application does not provide any additional information regarding this plasmid.

162. Plasmid #319-13, in turn, carries the coding region for the 187-amino acid precursor form of human fibroblast interferon, rather than that of the mature form of human fibroblast interferon having a total of 166 amino acids and unaccompanied by the human fibroblast interferon signal peptide. Furthermore: (a) #319-13 lacks the genetic element ("promoter") known to be required for transcription of a coding sequence to provide mRNA, (b) #319-13 lacks the genetic element ("Shine-Dalgarno sequence") identified as necessary for translation of mRNA into protein by a prokaryotic host; and (c) prokaryotic host cells, *e.g.*, *E. coli*, were not expected to be able to process a mammalian precursor protein faithfully so as to provide the naturally-occurring mature form thereof.

163. The specification of the Japan '931 application only discloses a coding sequence for the 166-amino acid mature form of human fibroblast interferon as the 498-base pair nucleotide sequence embedded within the 561-base pair sequence encoding the 187-amino acid precursor form of human fibroblast interferon.

164. The specification of the Japan '931 application only discloses the 166-amino acid sequence of the mature form of human fibroblast interferon as an amino acid sequence embedded within the 187-amino acid sequence of the precursor form of human fibroblast interferon. There is no explicit disclosure of the 166-amino acid sequence of mature human fibroblast interferon unaccompanied by its signal peptide.

165. As of March 19, 1980 the art did not provide a reasonable expectation that *E. coli* could process the 187-amino acid precursor form of human fibroblast interferon to provide the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by

the human fibroblast interferon signal peptide. In fact, the art suggested that *E. coli* would not process the 187-amino acid precursor form of human fibroblast interferon to provide the mature form of human fibroblast interferon.

166. The specification of the Japan '931 application does not disclose any method for removing the coding sequence of the human fibroblast interferon signal peptide to provide a coding sequence for the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

167. The specification of the Japan '931 application does not provide any example of a method for removing the coding sequence of the human fibroblast interferon signal peptide to provide a coding sequence for the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

168. The specification of the Japan '931 application does not provide any reference to a method for removing the coding sequence of the human fibroblast interferon signal peptide to provide a coding sequence for the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide. In fact, the specification of the Japan '931 application does not provide any reference to a method for removing the coding sequence of the signal peptide of any precursor protein, so as to provide a coding sequence for the mature form that protein.

169. The specification of the Japan '931 application does not describe an expression plasmid that would enable direct expression of the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

170. The specification of the Japan '931 application does not describe any method for constructing an expression plasmid that would enable direct expression of the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

171. The specification of the Japan '931 application does not describe expression of human fibroblast interferon in a microorganism.

172. The specification of the Japan '931 application does not describe a promoter that could be used to direct transcription of a DNA encoding the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

173. The specification of the Japan '931 application does not describe a "ribosome binding site" that could be used to enable translation of mRNA encoding the mature form of human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

174. Fig. 1 of the Japan '931 application (Exhibit 2013, at 17) indicates that the plasmid designated #319 lacks the HincII restriction site used in the tailoring experiments described in Goeddel (Exhibit 1012). This inference is consistent with Fig. 2 of Taniguchi *et al.* (Exhibit 1014), which indicates that the HincII site (GTYRAC, *i.e.*, GTCAAC), which is immediately upstream of the initiation codon of the precursor form of human fibroblast interferon, is missing from the DNA insert of plasmid #319 (Exhibit 1014, at 4004, Fig. 2, right col., last ¶ and at 4005, Fig. 4, bottom).

175. The plasmid designated #319-13 that is disclosed in the specification of the Japan '931 application does not have a promoter operatively associated with the coding sequence

encoding the 187 amino acid precursor form of human fibroblast interferon. In view of the fact that the DNA encoding the precursor form of the human fibroblast interferon was inserted into the EcoRI site of pBR322, and in view of the relative orientation of the ampicillin resistance gene and the tetracycline resistance genes of pBR322, it is unlikely that the DNA insert of #319-13 would be transcribed by either the amp^r or the tet^r promoter (Sutcliffe, "Complete nucleotide sequence of the *Escherichia coli* plasmid pBR322," Cold Spring Harb Symp Quant Biol., 43 Pt 1:77-90 (1979) (Exhibit 1032), at 84, Fig. 4; Ex. 2013 (p. 9, second ¶ through the first full ¶ at p. 10) indicating that the cDNA cloned by Sugano was inserted into the EcoRI site of the plasmid cloning vector pBR322).

176. The plasmid designated #319-13 that is disclosed in the specification of the Japan '931 application does not have a ribosome binding site that has been operatively associated with the coding sequence encoding the 187-amino acid precursor form of human fibroblast interferon.

177. Although the plasmid designated #319-13, which is disclosed in the specification of the Japan '931 application, does encode the 187-amino acid sequence of the precursor form of human fibroblast interferon, that plasmid would not be expected to direct expression of either the 187-amino acid precursor form of human fibroblast interferon or the 166-amino acid mature form of human fibroblast interferon

178. As noted above, to successfully produce a mature human fibroblast interferon free of its signal peptide by "direct expression" in bacteria, two key elements must be achieved: (a) DNA encoding the precursor form must be "tailored," or modified, to remove those codons encoding the entire signal peptide and (b) the ATG corresponding to the first amino acid of the mature form must be properly positioned with respect to bacterial expression control elements.

179. Nowhere in the specification of the Japan '931 application does Sugano describe a method for tailoring the human fibroblast interferon gene, which encodes the amino acids of the signal peptide, to remove those codons encoding the entire signal peptide. Nor does Sugano's specification describe a method for making an expression vector in which the ATG corresponding to the first amino acid of the mature form is properly positioned with respect to bacterial expression control elements such that it is capable of directing expression of a DNA encoding a mature human fibroblast interferon having a total of 166 amino acids and free of the fibroblast interferon signal peptide. In fact the specification of the Japan '931 application does not describe a method for the tailoring or the expression of any coding sequence.

180. In view of the information provided in the specification of the Japan '931 application, one of ordinary skill in the art would not have believed that, as of March 19, 1980, Sugano possessed a DNA encoding mature human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

181. In view of the information provided in the specification of the Japan '931 application, one of ordinary skill in the art would not have believed that, as of March 19, 1980, Sugano possessed a method for making a DNA encoding mature human fibroblast interferon having a total of 166 amino acids unaccompanied by the human fibroblast interferon signal peptide.

In signing this Declaration, I understand that the Declaration will be filed as evidence in a contested case before the Board of Patent Appeals and Interferences of the United States Patent and Trademark Office. I acknowledge that I may be subject to cross-examination in the case and that cross-examination will take place within the United States. If cross examination is required

of me, I will appear for cross-examination within the United States during the time allotted for-cross examination.

I declare further that all statements made herein of my own knowledge are true and that all statements made on information, belief and/or opinion are believed to be true and/or opinion are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Dated: 05/03/07

Rik Derynck, Ph.D.

Exhibits

Exhibit 1009	Taniguchi *et al.*, "Expression of the human fibroblast interferon gene in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 77: 5230-5233 (1980)
Exhibit 1010	Derynck *et al.*, "Isolation and structure of a human fibroblast interferon gene," Nature 285: 542-547 (1980)
Exhibit 1011	Derynck *et al.*, "Expression of human fibroblast interferon gene in *Escherichia coli*," Nature 287: 193-197 (1980)
Exhibit 1012	Goeddel, "Synthesis of human fibroblast interferon by *E. coli*," Nucl. Acids Res. 8: 4057-4074 (1980)
Exhibit 1013	Taniguchi *et al.*, "The nucleotide sequence of human fibroblast interferon cDNA," Gene 10: 11-15 (1980)
Exhibit 1014	Taniguchi *et al.*, "Molecular cloning of human fibroblast interferon cDNA," Proc. Nat'l Acad. Sci. USA 77: 4003-4006 (1980)
Exhibit 1016	Chapter 5: "Antigen-Antibody Interactions," PRINCIPLES OF IMMUNOLOGY, 2nd Ed., MacMillan Publishing Co., New York (1979), pp. 65-79
Exhibit 1026	Declaration of Jordan U. Gutterman, M.D., dated January 24, 2007
Exhibit 1028	Curriculum Vitae of Rik Derynck, Ph.D.
Exhibit 1029	Remaut *et al.*, "Inducible high level synthesis of mature human fibroblast interferon in *Escherichia coli*," Nucl. Acids Res. 11: 4677-4688 (1983)
Exhibit 1032	Sutcliffe, "Complete Nucleotide Sequence of the *Escherichia coli* Plasmid pBR322," Cold Spring Harb Symp Quant Biol., 43 Pt 1:77-90 (1979)
Exhibit 1036	Knight, "Interferon: Purification and initial characterization from human diploid cells," Proc. Nat'l Acad. Sci. PNAS 73: 520-523 (1976)
Exhibit 1037	Knight, "Human fibrboblast interferon: Amino acid sequence analysis and amino terminal amino acid sequence," Science 207: 525-526 (1980)
Exhibit 1043	Shine and Dalgarno, "Determination of cistron specificity in bacterial ribosomes," Nature 254: 34-38 (1975)
Exhibit 1057	Goeddel *et al.*, "Direct expression in *Escherichia coli* of a DNA sequence coding for human growth hormone," Nature 281: 544-548 (1979)
Exhibit 1058	Goeddel *et al.*, U.S. Patent No. 4,342,832, "Method of constructing a replicable cloning vehicle having quasi-synthetic genes," issued August 3,

1982

Exhibit 1067 Houghton *et al.*, "The absence of introns within a human fibroblast interferon gene," Nucl. Acids Res. 9: 247-266 (1981)

Exhibit 1068 Taniguchi *et al.*, "Expression of the Human Interferon-β_1 Gene in Heterologous Host Cells," *IN* "Proceedings of the IVth International Symposium on Genetics of Industrial Microorganisms (1982)" (Y. Ikeda & T. Beppu, eds.; Kobansha Ltd., Tokyo; 1983)

Exhibit 1078 UK Patent Application GB 2071671

Exhibit 1109 Steitz, "RNA·RNA Interactions During Polypeptide Chain Initiation," in RIBOSOMES Structure, Function and Genetics (Proceedings of the 9th. Steenbock Symposium Held at the University of Wisconsin, Madison, July 5-8, 1979; Ed. Chambliss *et al.*, University Park Press, Baltimore)

Exhibit 1110 Hautala *et al.*, "Increased expression of a eukaryotic gene in *Escherichia coli* through stabilization of its messenger RNA," Proc. Nat'l Acad. Sci. USA 76: 5574-5578 (1979)

Exhibit 1111 Goldberg, "Degradation of abnormal proteins in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 69: 422-426 (1972)

Exhibit 1112 Simon *et al.*, "Stabilization of proteins by a bacteriophage T4 gene cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 80: 2059-2062 (1983)

Exhibit 1113 Uren, "The recovery of genetically engineered mammalian proteins," Am. Biotechnol. Lab. 2: 51-54 (1983)

Exhibit 1114 Epstein, "The genetic control of tertiary protein structure: Studies with model systems," Cold Spring Harbor Symp. Quant. Biol. 28: 439-449 (1963)

Exhibit 1115 Fiers *et al.*, "Complete nucleotide sequence of bacteriophage MS2 RNA: Primary and secondary structure of the replicase gene," Nature 260: 500-507 (1976)

Exhibit 1116 Maxam and Gilbert, "A new method for sequencing DNA," Proc. Nat'l Acad. Sci. USA 74: 560-564 (1977)

Exhibit 1117 Gray *et al.*, "Extracellular nucleases of Pseudomonas BAL 31. I Characterization of single strand-specific deoxyriboendonuclease and double-strand deoxyriboexonuclese activities," Nucl. Acids Res. 2: 1459-1492 (1975)

Exhibit 1118 Ptashne *et al.*, U.S. Patent No. 4,418,149, "Fused hybrid gene," issued November 29, 1983

Exhibit 1119	Itoh *et al.*, "Efficient expression in *Escherichia coli* of a mature and a modified human interferon-β_1," DNA 3: 157-165 (1984)
Exhibit 1120	Itoh *et al.*, U.S. Patent No. 4,686,191, "Recombinant plasmid containing human interferon-beta gene," issued June 11, 1987
Exhibit 1121	European Patent Application No. 0064681 A2
Exhibit 1122	Rosenberg and Court, "Regulatory sequences involved in the promotion and termination of RNA transcription," Ann. Rev. Genetics 13: 319-353 (1979)
Exhibit 1123	Wei, "Isolation and comparisons of two molecular species of the BAL 31 nuclease from *Alteromonas espejiana* with distinct kinetic properties," J. Biol. Chem. 258: 13506-13512 (1983)
Exhibit 1124	Jackson *et al.*, "Biochemical method for inserting new genetic information into DNA of simian virus 40: Circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 69: 2904-2909 (1972)
Exhibit 1125	Lobban and Kaiser, "Enzymatic end-to end joining of DNA molecules" J. Mol. Biol. 78: 453-471 (1973)
Exhibit 1126	Roberts *et al.*, "A general method for maximizing the expression of a cloned gene," Proc. Nat'l Acad. Sci. USA 76: 760-764 (1979)
Exhibit 1127	Bendig *et al.*, "Deletion mutants of polyoma virus defining a nonessential region between the origin of replication and the initiation codon for early proteins," J. Virol. 32: 530-535 (1979)
Exhibit 1128	Shenk *et al.*, "Biochemical method for mapping mutational alterations in DNA with S1 nuclease: The location of deletions and temperature-sensitive mutations in simian virus 40," Proc. Nat'l Acad. Sci. USA 72: 989-993 (1975)
Exhibit 1129	Grantham *et al.*, "Codon catalog usage and the genome hypothesis," Nucl. Acids Res. 8: r49-r62 (1980)
Exhibit 1130	U.S. patent application Serial No. 06/131,152, which was filed March 17, 1980
Exhibit 1131	Weissmann, "The Cloning of Interferon and Other Mistakes," IN INTERFERON 3, Academic Press Inc. (New York, NY), 1981, pp. 101-134
Exhibit 2004	Ptashne *et al.*, U.S. Patent No. 4,332,892, "Protein synthesis," issued September 22, 1982
Exhibit 2013	English translation of Japanese Patent Application No. 033931/80

Exhibit 2059	Taniguchi *et al.*, "Construction and identification of a bacterial plasmid containing the human fibroblast interferon sequence," Proc. Jpn. Acad. 55: 464-469 (1979)
Exhibit 2067	Curtiss, "Genetic manipulation of microorganisms: Potential benefits and biohazards," Ann. Rev. Microbiol. 30: 507-533 (1976)
Exhibit 2068	Wu *et al.*, "Synthetic oligodeoxynucleotides for analyses of DNA structure and function," Prog. Nucl. Acid Res. Mol. Biol. 21: 101-141 (1978)
Exhibit 2069	Erlich *et al.*, "A sensitive radioimmunoassay for detecting products translated from cloned DNA fragments," Cell 13: 681-689 (1978)
Exhibit 2078	Ratzkin and Carbon, "Functional expression of cloned yeast DNA in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 74: 487-491 (1977)
Exhibit 2079	Vapnek *et al.*, "Expression in *Escherichia coli* K-12 of the structural gene for catabolic dehydroquinase of *Neurospora crassa*," Proc. Nat'l Acad. Sci. USA 74: 3508-3512 (1977)
Exhibit 2080	Itakura *et al.*, "Expression in *Escherichia coli* of a chemically synthesized gene for the hormone somatostatin," Science 198: 1056-1063 (1977)
Exhibit 2081	Villa-Komaroff *et al.*, "A bacterial clone synthesizing proinsulin," Proc. Nat'l Acad. Sci. USA 75: 3727-3731 (1978)
Exhibit 2082	Mercereau-Puijalon *et al.*, "Synthesis of an ovalbumin-like protein by *Escherichia coli* K12 harbouring a recombinant plasmid," Nature, 275: 505-510 (1978)
Exhibit 2083	Chang *et al.*, "Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dihydrofolate reductase," Nature 275: 617-624 (1978)
Exhibit 2084	Fraser and Bruce, "Chicken ovalbumin is synthesized and secreted by *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 75: 5936-5940 (1978)
Exhibit 2085	Seeburg *et al.*, "Synthesis of growth hormone by bacteria," Nature 276: 795-798 (1978)
Exhibit 2086	Goeddel *et al.*, "Expression in *Escherichia coli* of chemically synthesized genes for human insulin," Proc. Nat'l Acad. Sci. USA 76: 106-110 (1979)
Exhibit 2087	Schell *et al.*, "Cloning and expression of the yeast galactokinase gene in an *Escherichia coli* plasmid," Gene, 5: 291-303 (1979)
Exhibit 2088	Bach *et al.*, "Evidence for transcriptional regulation of orotidine-5'-phosphate decarboxylase in yeast by hybridization of mRNA to the yeast structural gene

cloned in *Escherichia coli*," Proc. Nat'l Acad. Sci. 76: 386-390 (1979)

Exhibit 2089 Burrell *et al.*, "Expression in *Escherichia coli* of hepatitis B virus DNA sequences cloned in plasmid pBR322," Nature 279: 43-47 (1979)

Exhibit 2090 Martial *et al.*, "Human growth hormone: Complementary DNA cloning and expression in bacteria," Science 205: 602-607 (1979)

Exhibit 2092 Wilson *et al.*, "Detection of proteins like human gamma and beta globins in *Escherichia coli* carrying recombinant DNA plasmids," Proc. Nat'l Acad. Sci. USA 76: 5631-5635 (1979)

Exhibit 2093 Roberts *et al.*, "Synthesis of simian virus 40 t antigen in *Escherichia coli*," Proc. Nat'l Acad. Sci. USA 76: 5596-5600 (1979)

Exhibit 2094 Emtage *et al.*, "Influenza antigenic determinants are expressed from haemagglutinin genes cloned in *Escherichia coli*," Nature 283: 171-174 (1980)

Exhibit 2095 Backman *et al.*, "Maximizing gene expression on a plasmid using recombination *in vitro*," Cell 13: 65-71 (1978)

Exhibit 2126 Goeddel *et al.*, "Human leukocyte interferon produced by *E. coli* is biologically active," *Nature* 287: 411-416 (1980)

Exhibit 2132 Guarente *et al.*, "Improved methods for maximizing expression of a cloned gene: A bacterium that synthesizes rabbit β-globin," Cell 20: 543-553 (1980)